Karl Popp
Werner Schiehlen (Eds.)

System Dynamics and Long-Term Behaviour of Railway Vehicles, Track and Subgrade

Springer

Lecture Notes in Applied Mechanics

Volume 6

Series Editor
Prof. Dr.-Ing. Friedrich Pfeiffer

Springer-Verlag Berlin Heidelberg GmbH

Engineering ONLINE LIBRARY

http://www.springer.de/engine/

System Dynamics and Long-Term Behaviour of Railway Vehicles, Track and Subgrade

Karl Popp and Werner Schiehlen (Eds.)

Springer

Professor
Dr.-Ing. KARL POPP
Universität Hannover
Institut für Mechanik
Appelstraße 11
30167 Hannover
GERMANY
e-mail: popp@ifm.uni-hannover.de

Professor
Dr.-Ing. WERNER SCHIEHLEN
Universität Stuttgart
Institut B für Mechanik
Pfaffenwaldring 9
70550 Stuttgart
GERMANY
e-mail: wos@mechb.uni-stuttgart.de

Preface

High speed ground traffic is strongly competitive to air traffic in densely populated areas like Europe or Japan. During the last decades, completely new technologies for high speed railway vehicles were developed. The primarily goals have been to achieve travelling speeds from 250 to 350 km/h and to guarantee safety of the passengers. In Germany these goals were met by the ICE train service starting in 1991. However, certain phenomena were not considered thoroughly during the development of high speed systems: track response, rail and wheel corrugation and noise generation. It turned out that there was a challenge for interdisciplinary research bringing together civil and mechanical engineers dealing with railway problems.

In 1996 the Deutsche Forschungsgemeinschaft (German Research Council) accepted the proposal initiated by Professors K. Knothe (Berlin), P. Meinke (Stuttgart) and K. Popp (Hannover) for a Priority Programme devoted to *System Dynamics and Long-Term Behaviour of Vehicle, Track and Subgrade.*

The aims of the Priority Programme were to contribute to a better understanding of the dynamic interaction of vehicle and track and the long-term behaviour of the components of the entire system. So far, the approach has been to increase speed, traction and axle-load by optimizing sub-systems. However, the entire system has revealed new limitations: settlement and destruction of the ballast and the subgrade lead to deterioration of the track; irregular wear of the wheels cause an increase in overall load and deterioration of passengers' comfort; damage of the running surfaces of the rail and the wheel becomes more frequent.

The research was based on four main issues. Non-linear behaviour of the ballast was investigated experimentally in laboratories and simulated by new material-laws and multi-particle computer models. The coupling of the track-model and the subgrade was defined using various methods. Phenomena occurring in the contact area of wheel and rail were numerically and experimentally analyzed, also by using a virtual measurement device on the moving train. Dynamic models of the entire system were thought to explain the phenomenon of polygonalization of the wheels. It was expected that the interdisciplinary and fundamental research will result in physical explanations and subsequently will lead to engineering developments which will allow long-lasting solutions.

All the aforementioned engineering problems have been investigated in the Priority Programme SPP 1015 with research activities financially supported by the Deutsche Forschungsgemeinschaft (DFG) from 1996 to 2002. The research comprised 22 interconnected projects conducted at 16 German universities and research institutions. The research activities were organized in the following sections:

- Vehicle Dynamics (coordinator: P. Meinke, Stuttgart)
  - Wavy wear pattern on the tread of railway wheels (E. Brommundt, Braunschweig)
  - Distributed numerical calculations of wear in the wheel-rail contact (K. Frischmuth, Rostock)
  - Modeling and simulation of the mid-frequency behaviour of an elastic bogie (K. Popp, Hannover)
  - Rotor dynamics and irregular wear of elastic wheelsets (P. Meinke, Stuttgart)
  - System dynamics of railcars with radial- and lateralelastic wheels (W. Schiehlen, Stuttgart)
- Contact, Friction, Wear (coordinator: B. Zastrau, Dresden)
  - Model-based validation within the rail-wheel-subgrade modeling (D. Söffker, Duisburg)
  - Friction and wear of tractive rolling contacts (L. Deters, Magdeburg)
  - Experimental analysis of the cyclic deformation and damage behavior of characteristic wheel and rail steels (K.-H. Lang, Karlsruhe)
  - On the numerical analysis of the wheel-rail system in rolling contact (B. Zastrau, Dresden)
- Track Dynamics (coordinator: K. Knothe, Berlin)
  - Monitoring the dynamics of railway tracks by means of the Karhunen-Loève-Transformation (E. Kreuzer, Hamburg-Harburg)
  - Measurements and modelling of resilient rubber rail-pads (K. Knothe, Berlin)
  - Combined modelling of discretely supported track models and subgrade models - Vertical and lateral dynamics (K. Knothe, Berlin)
  - The dynamics of railway track and subgrade with respect to deteriorated sleeper support (R. Lammering, Hamburg)
  - Model-based investigation of the dynamic behaviour of railway ballast (K. Popp, Hannover)
- Subgrade Dynamics (coordinator: G. Gudehus, Karlsruhe)
  - Experimental and numerical investigations on the track stability (R. Katzenbach, Darmstadt)
  - Simulation of the dynamic behavior of bedding-foundation-soil in the time domain (O. v. Estorff, Hamburg-Harburg)
  - Numerical model and laboratory tests on settlement of ballast track (G. Gudehus, Karlsruhe)
  - Experimental investigation and numerical modelling of soils and ballast under cyclic and dynamic loading (H.-G. Kempfert, Kassel)

- Rigid body dynamics of railway ballast (T. Pöschel, Berlin)
- Track settlement due to cyclic loading with low minimum pressure and vibrations (W. Ruecker, Berlin)
- 3D-simulation of dynamic interaction between track and layered subground (S. Savidis, Berlin)
- Dynamic behavior of railway track systems analyzed in frequency domain (G. Schmid, Bochum)

This book contains the scientific results of the Priority Programme as presented at the Concluding Colloquium held at the University of Stuttgart, Germany, March 13 - 15, 2002. Additionally, the earlier concluded projects are also included in this volume as well as an introduction with review and outlook, and a benchmark study.

The following experts in the field were invited to join the Colloquium to present survey lectures on related research topics:

- H. P. Lang: Experience from high-speed traffic
- R. D. Fröhling: Track and subgrade loading
- P. Pointner: The wheel-rail-contact and its effect on wear and rolling-contact-fatigue-cracks
- R. M. Goodall: Active suspension technology and its impact upon vehicle/track interaction

Three of their contributions have been included in this book, too. In summary, this book represents a final scientific report of the DFG Priority Programme SPP 1015.

Finally, we would like to thank the Deutsche Forschungsgemeinschaft (DFG) for the financial support and, in particular, to its representative, Dr.-Ing. Jürgen Hoefeld, for his continuing interest in the Programme. We acknowledge the excellent work of the section coordinators, Professors G. Gudehus, K. Knothe, P. Meinke and B. Zastrau who arranged for the fruitful cooperation within the Programme. And last but not least we thank our coworkers Dipl.-Ing. T. Meinders and Dipl.-Ing. H. Kruse for their most valuable help during the preparation of the Concluding Colloquium and the preparation of this Proceedings Volume.

Hannover, May 2002
Karl Popp

Stuttgart, May 2002
Werner Schiehlen

Contents

SUBGRADE DYNAMICS 315

Numerical Model and Laboratory Tests on Settlement of Ballast Track 317
Sven Augustin, Gerd Gudehus, Gerhard Huber,
Andreas Schünemann

Track Settlement Due to Cyclic Loading with Low Minimum Pressure and Vibrations 337
Matthias Baessler, Werner Ruecker

The DFG Priority Programme
System Dynamics and Long-Term Behaviour
of Vehicle, Track and Subgrade

Klaus Knothe[1] and Karl Popp[2]

[1] Technische Universität Berlin, Institut für Luft- und Raumfahrt,
   Marchstr. 12, D-10587 Berlin, Germany
[2] Universität Hannover, Institut für Mechanik,
   Appelstr. 11, D-30167 Hannover, Germany

Abstract. This paper deals with the motivation and the aims of the DFG Priority
Programme 1015 *System Dynamics and Long-Term Behaviour of Vehicle, Track
and Subgrade* as well as its results. The authors try to answer the following three
questions: i) Where did we start from, ii) what have we achieved and iii) what still
needs to be done. The critical assessment of the results intends to point out topics
of further research.

1   Introduction

This introductory paper gives an outline of the DFG Priority Programme on
System Dynamics and Long-Term Behaviour of Vehicle, Track and Subgrade.
It goes back to the roots, shows the motivation and aims and tries to give a
critical evaluation of the results achieved. The paper is based on the state-
ments given during the opening lecture by K. Knothe and the closing remarks
by K. Popp, presented at the Concluding Colloquium held in Stuttgart from
March 13 to 15, 2002.

The DFG Priority Programme started in 1996. More than 20 projects
were supported by Deutsche Forschungsgemeinschaft with a budget of more
than Euro 6 million. At the end of the Priority Programme it is tried to
answer the following questions:

- Where did we start from?
- What have we achieved?
- What still needs to be done?

The critical evaluation of the results gained so far on the one hand and the
discussions of the requirements of today's railway systems on the other hand
shall give an account of the achievements of the programme. Based on this
evaluation, further topics of research can be pointed out.

Fig. 1. Deformation of an SNCF track (1955) after a world record ride, probably due to unstable hunting motion of the locomotive (from *La Vie du Rail*, 1981)

2   Where did we start from?

2.1   Preliminary Comments

Talking about the starting point of this DFG Priority Programme, it ist not enough to go back to the first discussions which led to this Programme – one has to go back further. Although it would be interesting to consider 150 years of railway research, this, of course, would be beyond the scope of this paper. Therefore, we shall start in 1975.

2.2   Limits of the Wheel/Rail System (1975-1985)

Reasons for the investigations. At the beginning of the 1970s the companies of the so-called New Technologies decided to build up an alternative to the wheel/rail system, the Maglev train. This project was supported by the Federal Ministry of Research and Technology in Germany. The result

of this initiative is the Transrapid. On the other hand, the traditional German rolling stock industry (especially Krupp company) was afraid that it could lose market shares and, therefore, also claimed financial support for a research and development programme on wheel/rail problems. One of the main questions was: Are there any physical limits to the wheel/rail system? In order to answer this question, a first research and development project was supported by the Federal Ministry of Research and Technology in 1974. Participants were the West German railway industry (among them the companies Krupp and Thyssen) and the Deutsche Bundesbahn. Some years later additional projects were initiated by the companies MAN and MBB together with universities and research institutes.

It should be mentioned that in Japan and France high speed wheel/rail technologies were developed much earlier. In Japan research projects started immediately after 1945, in France five or six years later.

Aims of the investigations. One of the aims of the investigations was to find out whether there is a critical maximum speed of the wheel/rail system. It was already known that an optimum speed would be between 300 and 350 km/h due to energy problems. Therefore, the aim was specified:

- Would it be possible to run wheel/rail systems with speeds up to 350 km/h?

- What has to be done so that German industry would be able to build appropriate locomotives and passenger cars?

An example. Let us consider the stability problem as an example of the research work done between 1975 and 1985.

The fact that railway vehicles may become unstable is known from the work of Carter in 1916 [1]. During a test run with two newly developed high speed locomotives in France in 1955 serious problems were experienced with an unstable running locomotive. Both rails were thrown to the left and to the right by one of the locomotives, and the wire of the power supply was pulled down by the pantograph (Fig. 1). These events were not made public until 1981 [5].

Nowadays a linear stability analysis is state-of-the-art for all newly designed vehicles (Fig. 2) [3]. Root loci indicate vehicle speed for which positive real parts of the eigenvalues occur so that instabilities have to be expected.

Even nonlinear stability analysis is common as shown in Fig. 3. For a vehicle speed of 72 m/sec the wheelset performs a limit cycle with an amplitude of lateral displacement of more than 6 mm. The vehicle speed has to be reduced to 58 m/sec to let the wheelsets return to the central position. This is a typical nonlinear phenomenon. More detailed investigations have been performed by True and his group [7,4,8].
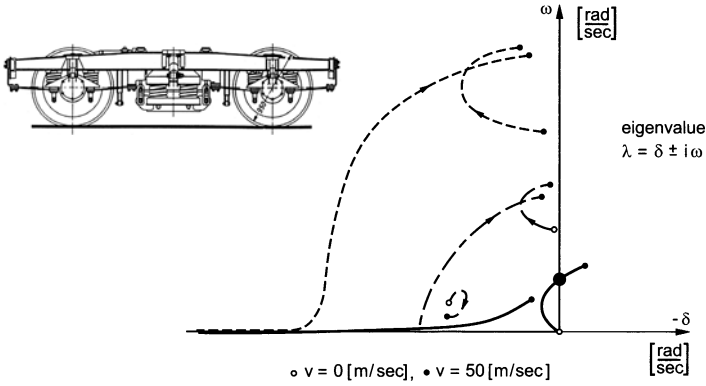
Fig. 2. Linear stability analysis of a bogie with secondary suspension
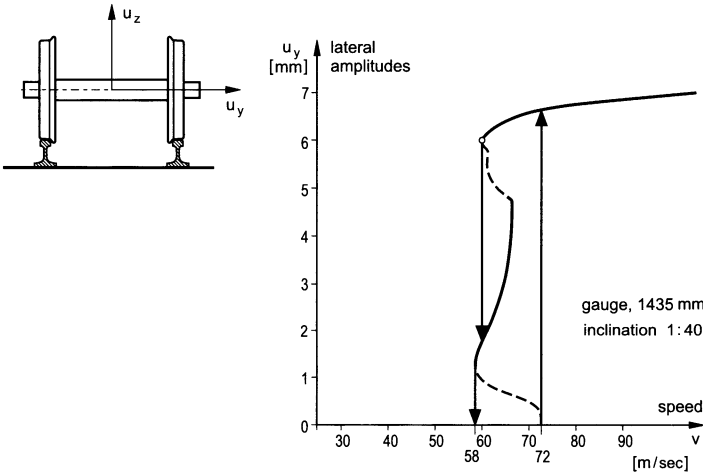
Fig. 3. Nonlinear limit cycle analysis of a conventional bogie, taken from [2]

A similar nonlinear instability phenomenon has been observed also by MAN [6] on the roller-rig in Freimann near Munich. If the vehicle speed was increased to 180 km/h a limit cycle with an amplitude of more than 7 mm occurred. One had to reduce the vehicle speed to less than 130 km/h until the original stable situation was obtained again.

Results of the investigations between 1975 and 1985. The results of the German Research and Development Programme between 1975 and 1985 are well known.

- The knowledge of the German railway industry about the design of high speed railway vehicles rapidly increased.
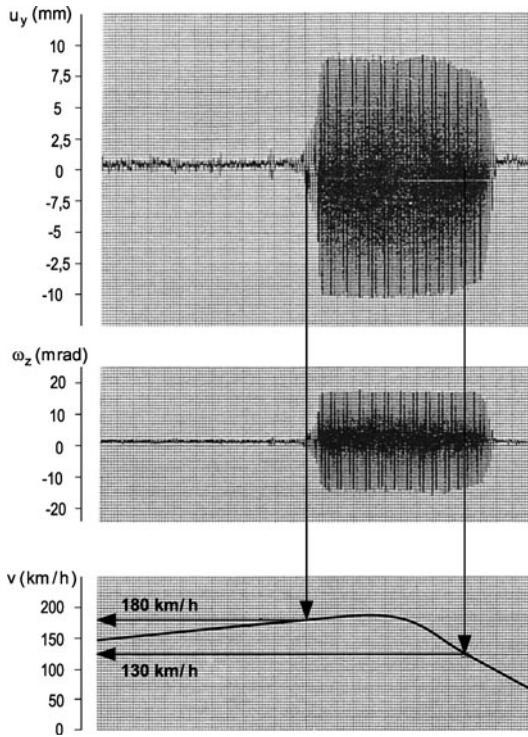
Fig. 4. Experimental investigations of limit cycles, taken from [6]

- New simulation tools were introduced in industry and in railway companies. At least two computer codes (MEDYNA and SIMPACK) were developed in Germany and are now available for all tasks in railway vehicle dynamics.
- And last but not least: The Intercity Experimental train was developed as a prototype for the later ICE train.

2.3   New Problems (1985-1994)

At the end of the 1980s and at the beginning of the 1990s a remarkable increase in vehicle speed has been obtained and high speed traffic has become common in Germany. However, several new problems occurred with vehicles and track. Though most of them have been observed earlier, they now occurred much more frequently.

Irregular ballast settlement and ballast deterioration. Irregular ballast settlement is one of the failures of railway tracks which cause high maintenance costs. Not only in Germany but also in France on the line of the TGV Atlantique and in Japan irregular ballast settlements, sometimes together with damage phenomena like the pulverization of ballast (white

spots), increased with increasing vehicle speed. For example it was not clear to the French railway SNCF whether the line to Marseille should be built as a line with ballasted track.

Rail corrugation. Short wavelength rail corrugation is a damage phenomenon which has been known fore 100 years. With increasing vehicle speed, however, rail corrugation increased signifficantly and again resulted in high maintenance costs, not only in Germany but also in Japan.

Out-of-round wheels. The phenomenon of out-of-round or polygonalized wheels of the newly introduced ICE 1 train was a very undesirable effect which first resulted in considerable comfort reduction for passengers. Later on the polygonalisation phenomenon of a rubber supported wheel rim was one of the causes of the Eschede accident in 1998.

Increasing noise. Rail corrugation immediately leads to severe noise problems, if no remedies like rail grinding are applied.

2.4   Explanation and Counter-Measures

Explanation: What has gone wrong? At the end oft the 1980s, some scientists became convinced that there was a causal link between increased high-speed traffic on the one hand and increasing negative side-effects on the other.

In Germany until approximately 1980 only moderate technical modifications of the wheel/rail system had been introduced. A kind of *evolutionary optimum* had been reached. The maximum vehicle speed remained in a moderate range.

In order to be able to run railway vehicles at much higher speed, several modifications had to be introduced. Three of them shall be mentioned:

- First of all the traction had to be increased. Thus, controlled traction units were built.
- New bogies were built for ICE passenger cars or existing bogies were modified. As an example, the wheelsets were equipped with four disc brakes instead of two, so the unsprung mass increased.
- Additionally, the tracks of high speed lines were made much stiffer than ordinary tracks.

The main criterion regarding optimization had been vehicle speed, which resulted in a *subsystem optimization*.

Basic demands for new developments. To overcome the problems generated by the subsystem optimization, three demands had to be fulfilled:

- Vehicle dynamics had to be substituted by vehicle/track dynamics, i.e. by real system dynamics.

- The restriction to the low frequency range had to be abandoned. In addition, the medium frequency range had to be included in the dynamic investigations.
- The analysis of the short time system dynamics had to be augmented by the analysis of the long-term behaviour of different vehicle and track components.

These three demands, which were formulated at the beginning of the 1990s, were the basic ideas of the DFG Priority Programme.

Research as counter reaction. Based on these ideas already at the beginning of the 1990s J.-M. Lipsius from Thyssen, Kassel, and P. Meinke, at that time with MAN New Technology, Munich, together with others tried to convince the Federal Ministry of Research and Technology to establish a new research Programme on long-term phenomena of the wheel/rail system. They were not successful, since the funds were needed for the Maglev train Transrapid.

More successful was a European initiative on ballasted tracks called EU-ROBALT. Starting in 1993, railway companies from Germany, France and the UK, research institutes and ERRI worked together in two projects funded in the BRITE-EURAM frame.

3   What has been achieved?

Before answering this second question a look at the original proposal for the DFG Priority Programme shows the intentions. This proposal was intensively discussed and fixed by K. Knothe, P. Meinke and K. Popp during a German-Polish Workshop in Wierzba in July, 1993.

3.1   The Project Idea of the DFG Priority Programme

Those who had initiated the DFG Priority Programme had a clear idea about what should be integrated in the Programme and what should be excluded.

The components of vehicle and track which should be considered are shown in Fig. 5. They start from the bogie frame and go down to the track formation and the subgrade.

The system dynamics should have its emphasis in the medium frequency range, i.e. between approximately 40 and 400 Hz.

Of course, the wheel/rail contact as centre of interaction was considered to be extremely important.

Other types of interaction like pantograph/wire or aerodynamic interaction were consciously excluded, as well as all acoustic problems and ground-borne vibrations, cf. Fig. 6.
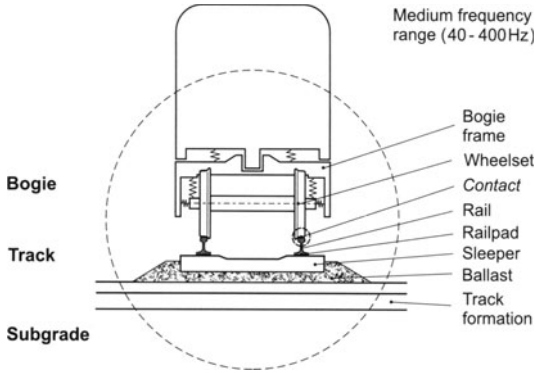
Fig. 5. Components to be considered in the DFG Priority Programme
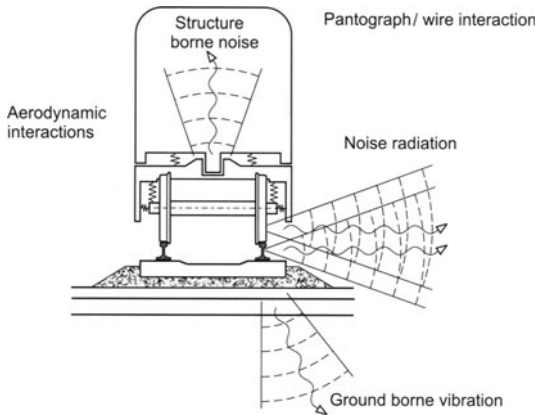
Fig. 6. Components to be excluded from the DFG Priority Programme
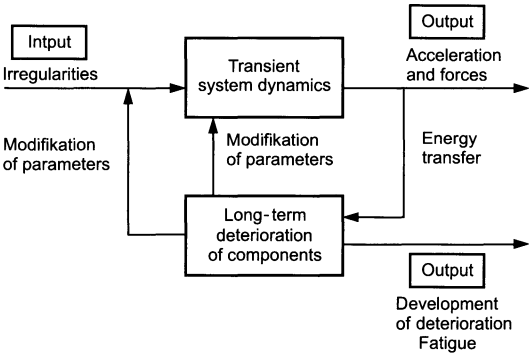
Fig. 7. System dynamics and long-term behaviour

Already the title of the Priority Programme *System Dynamics and Long-Term Behaviour of Vehicle, Track and Subgrade* indicated that *System Dynamics* and *Long-Term Behaviour* were two equally important aspects.

Long-term behaviour is often integrated into a feed-back loop. It is worth to call in mind the diagram of Fig. 7, where system dynamics and long-
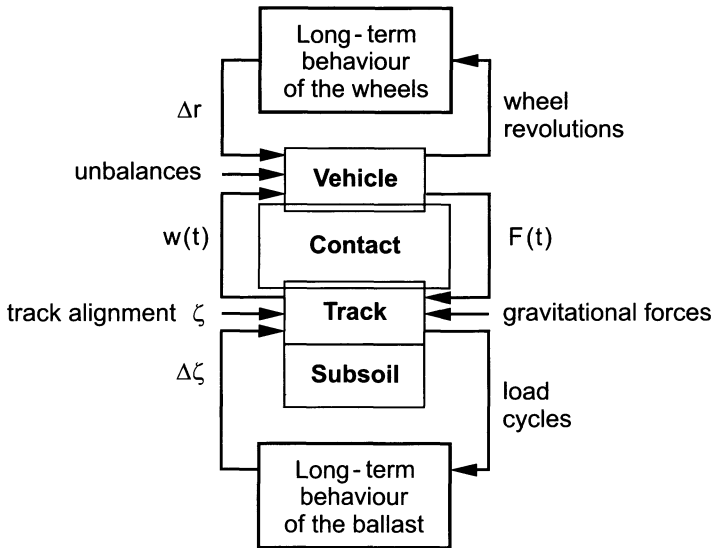
Fig. 8. Structure of the wheel/rail system

term behaviour are brought together to form this loop. The same type of diagram is assumed to be valid for different phenomena of deterioration, like rail corrugation, polygonalisation of wheels or settlement and degradation of ballast or subgrade, see also Fig. 8. In this context, it is essential to determine the evolution laws of the long-term behaviour and then to investigate the complete feed-back loop.

3.2   About Structure and Management of the DFG Priority Programme

The DFG Priority Programme was officially approved in January 1996 and lasted three terms of two years each. It was a truly interdisciplinary Programme, where mechanical and civil engineers, physicists and mathematicians worked together. The programme comprised more than 20 projects which had been grouped into one of the following four sections:

- Vehicle (coordinator: P. Meinke, Stuttgart)
- Contact, Friction, Wear (coordinator: B. Zastrau, Dresden)
- Track (coordinator: K. Knothe, Berlin)
- Subgrade (coordinator: G. Gudehus, Karlsruhe)

The sections were coordinated separately and each section organized about two meetings per year. Meetings of all members of the Priority Programme were called in by DFG:

- Kick-off Meeting, Braunschweig, 22.02.1996
  (organized by E. Brommundt)

- Intermediate Colloquium, Stuttgart, 20./21.03.1997
  (organized by W. Schiehlen)
- Intermediate Colloquium, Hamburg, 17./18.03.1999
  (organized by R. Lammering)
- Concluding Colloquium, Stuttgart, 13.-15.03.2002
  (organized by W. Schiehlen)

Moreover, the members of the DFG Priority Programme discussed their results in the framework of an international conference:

- EUROMECH Colloquium 409, Hannover, 6.-9.03.2000
  (organized by K. Popp, R. Bogacz and H. True)

One of the first problems to be solved within the sections was the definition of common data for the system parameters in order to compare research results gained in the different projects. As a consequence, the so-called DFG Reference Data Sets *Vehicle* and *Track* were created (see also *http://www.mechb.uni-stuttgart.de/forschung/railway/refdataA* and *http://ice.fb12.tu-berlin.de/gleisdaten*).

Later on, a benchmark data set *Track – Subsoil* was defined for a comparative study of different track-subsoil models. Results of this benchmark are given in an extra contribution inside this book. Both initiatives turned out to be very fruitful.

3.3   What has been achieved?

First, the different projects of the DFG Priority Programme shall be arranged in groups with similar subjects in order to illustrate where the emphasis was put on. The arrangement is done according to

- the vehicle and track components, which are mainly treated,
- the system dynamics on the one hand and the long-term behaviour on the other,
- the methods which are used, either simulation or experiment.

The result is illustrated in Fig. 9, where the number of projects is shown in brackets.

In the field of *system dynamics*, simulations are absolutely predominant.

- Three projects deal with linear or nonlinear simulations of the system behaviour. In all of them the subgrade is included and interactions between different sleepers through the subgrade are considered.
- In addition a group of four projects concentrates on the simulation of the dynamic behaviour of subgrade and ballast.
- The dynamic behaviour of wheelsets and bogie frame in the medium frequency range is considered in two projects.

Concerning the investigation of the *long-term behaviour* there are
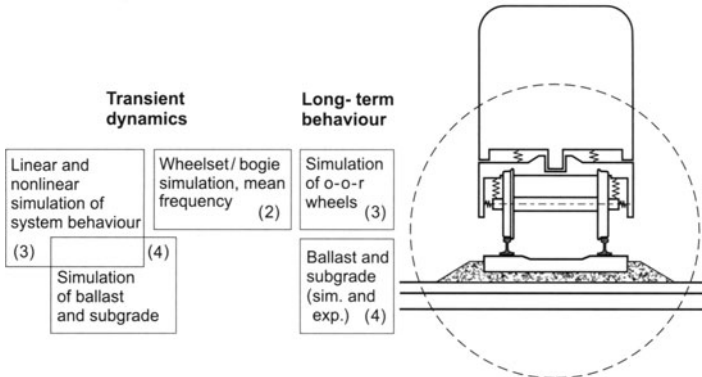
Fig. 9. Arrangement of projects with similar subjects (part I)

- three projects dealing with the simulation of out-of-round wheels and
- four projects dealing with the long-term behaviour of ballast and sub-grade, on the basis of experiments as well as simulations.

Some comments on these projects:

- Due to the activities in the DFG Priority Programme, linear and non-linear simulation of vehicle/track interaction in the medium frequency range is well understood.
- The same is true for the simulation of ballast and subgrade as an infinite, continuous or layered half-space with or without inclusions.
- Two or three projects on the system dynamics of wheelsets and bogie cannot cover all problems of these components in the medium frequency range. However, it is demonstrated that the tools required for such investigations are fully available.

The situation is not as clear concerning projects dealing with the long-term behaviour:

- Two hypotheses are used for the investigation of the polygonalisation of wheels. It is shown that with both hypotheses the development of out-of-round wheels can be explained qualitatively. Since there are more than two hypotheses, it would be worth bringing together all people from Germany or even from Europe working in the field of out-of-round wheels and to summarize different aspects of this problem.
- Settlement and deterioration of ballast and subgrade is a fascinating field. The work done in the DFG Priority Programme is a big step forward regarding the understanding of the phenomena. We assume that within five to ten years of research ballast settlement will be practically understood. In order to accelerate this development, a workshop on ballast deterioration problems would be helpful.
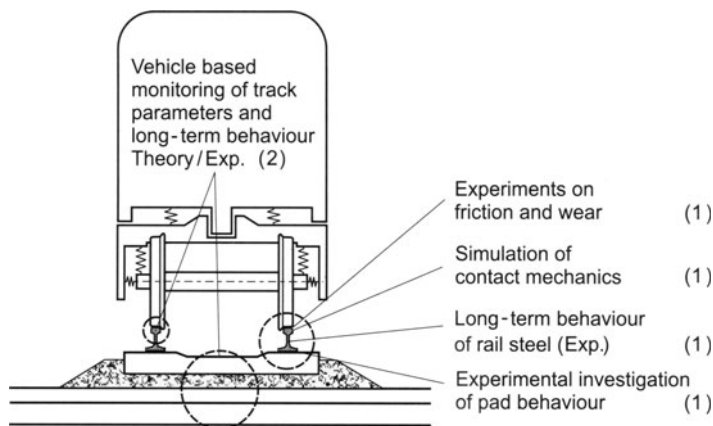
Fig. 10. Arrangement of projects with similar subjects (part II)

Problems which are only dealt with by one or two projects and which have not been mentioned so far are shown in Fig. 10.

- One project concentrates on the experimental investigation of friction and wear in rolling contact.
- In a second project a three-dimensional finite-element analysis of the rolling contact is performed.
- In the third project measurements are made investigating the long-term behaviour of rail steel under cyclic loading.
- And finally, the vertical dynamic behaviour of one type of rail pad is measured.

All these four projects provide valuable contributions. However, they cannot be considered as a final treatment of the problems. This is especially true for the wheel/rail contact which is the most important component of the wheel/rail system. In this context, it should also be stressed that rolling contact fatigue which now is considered to be the most critical failure of the rail surface was not even mentioned in the proposal for the DFG Priority Programme and, hence, was not investigated in any of the projects.

Finally, two projects remain which are both theoretical but should be applied in the future for vehicle based measurements. Unfortunately, one of these projects only lasted for four years and the other one only for two years. Vehicle based monitoring of long-term variations of contact and track parameters remains a challenging task.

4   What still needs to be done?

Simulation tools alone are never enough. What is necessary is an experimental validation of simulation programs for system dynamics of vehicle/track

interaction and for long-term behaviour of ballast and subgrade on a full-scale-level. This can only be performed with the support of railway operating companies like Deutsche Bahn AG in Germany.

Considering the last two projects mentioned above it is obvious that vehicle-based monitoring of track parameters and of long-term variations of the track is a task for the coming years. Again it must be ensured that Deutsche Bahn AG or another railway company will support corresponding measurements.

We are thoroughly convinced that fundamental aspects of the wheel/rail contact will be even more important in the future than they are today:

- A first problem which still is unsolved concerns in-situ measurement techniques for wheel/rail contact phenomena: *Measuring in the contact patch.*
- *Damage phenomena of the running surface* of wheel and rail continue to be of interest. There is a lack of knowledge of the material behaviour for extreme loading conditions in the wheel/rail contact. The same is true for other phenomena of long-term behaviour. In this context rolling contact fatigue has to be mentioned again.
- A wide field are *fundamental physical aspects of friction and wear in rolling contact.*
- A fourth challenging task is the so-called *conditioning of rail surfaces* either naturally by wheels running on the rail, or artificially.

To solve these problems, probably, other groups of scientists would be necessary than those which were partners in the present DFG Priority Programme.

Acknowledgements

References

1. Carter F.W. (1916) The electric locomotive. Proc. Inst. Civil Engn. 201:221–252, Discussion on pages 253–289
2. Gasch R., Moelle D., Knothe K. (1984) The effects of non-linearities on the limit-cycles of railway vehicles. In: The Dynamics of Vehicles on Roads and on Tracks, Proceedings of the 8th IAVSD-Symposium held at Cambridge, Mass./USA, 1983, Swets & Zeitlinger, Lisse/Amsterdam, 207–224

3. Gerdsmeier H., Budde U., Waldeck H. (1991) Lauftechnik für den Hochgeschwindigkeitsverkehr. ETR 40(10):649–656
4. Knudsen ., Feldberg R., True H. (1992) Bifurcations and chaos in a model of a rolling railway wheelset. Phil. Trans. R. Soc. Lond. A 338:455–469
5. Salin J. (1981) Regards en arrière. Revue Générale des Chemins de Fer 100:216–221
6. Schmidt A., Mauer L. (1982) Durchführung der Rollprüfstands-Versuche. Abschlußbericht zum BMFT-Forschungsvorhaben TV 7960 4. Bericht K 096 991 - EDS - 9. April 1982, MAN Neue Technologie München
7. True. H. (1992) Railway vehicle chaos and asymmetric hunting. In: The Dynamics of Vehicles on Roads and on Tracks, Proc. of the 12th IAVSD Symposium held at Lyon/France, 1991, Swets & Zeitlinger, Amsterdam/Lisse, 625–637
8. True H. (1994) Does a critical speed for railroad vehicles exist? In: Hawthorne K.L., Hill R.J. (Eds.) Proceedings of the 1994 ASME/IEEE Joint Railroad Conference held in Chicago, Illionos, 1994, 125–131

Part I:
Invited Lectures

Vehicle/Track Interaction Optimisation within Spoornet

Robert Fröhling

Spoornet, Railway Engineering, Private Bag X637, Pretoria, 0001, South Africa

Abstract. This paper gives a practical perspective on the optimisation of the interaction between vehicle and track within Spoornet over almost 40 years. After giving a historic overview of the development of the self-steering three-piece bogie, which plays an important part in vehicle/track interaction optimisation, a detailed account is given of the wheel/rail optimisation strategies followed on the iron ore and coal export lines. This is followed by a general description of the wheel profile development, measuring and monitoring actions, as well as a brief discussion on envisaged future developments aimed to further optimise the dynamic interaction between vehicle and track.

1   Introduction

The objective of vehicle/track interaction optimisation is to reduce forces or damaging mechanisms acting between these two systems. This can be achieved through track-friendly bogie suspension designs that have a good tracking ability and minimised dynamic vertical track loading. A further important element in this optimisation process is to control wheel and rail profile wear while minimising the contact stress between the wheel and the rail.

These elements of vehicle/track optimisation are discussed in this paper by giving a historic overview of the development of the self-steering three-piece bogie in South Africa together with a detailed account of the wheel/rail optimisation strategies followed on the iron ore and coal export lines. A general description is also given of the wheel profile development together with relevant measurement and monitoring actions that have been implemented to manage this optimisation process. This is followed by a short description of differential track settlement, which is a function of the spatial variation of the track stiffness as well as the dynamic interaction between the vehicle and the track. Finally a brief overview on envisaged future developments to further optimise the dynamic interaction between vehicle and track is given.

2   Wagon Bogie Development

In this section a brief overview of the development of the self-steering three-piece bogie is given. Further detail can be found in numerous international papers [1–16].

Due to South Africa's large deposits of iron ore in the Sishen area, some 800km from the coast, plans were made in the 1960s for large-scale iron ore export. Initially the harbour in Port Elizabeth was selected and the iron ore was to be transported on the existing line with existing rolling stock. The available wagons had a payload of 51ton and were equipped with three-piece Sarcast bogies with plain bearings as shown in Fig. 1. Due to a 31.4km long section of track with an almost continuous down grade of 1/66, train handling proved to be very difficult using the existing vacuum brake system. Hence, the length of the trains was limited to 35 wagons.

A subsequent economic study showed that the operation could be improved considerably if the train mass could be increased. Hence, wagons with an axleload of 18.5ton were built. In 1970 these wagons were fitted with three-piece roller bearing Spoorbarber bogies as shown in Fig. 4. Using these wagons, 50 wagon long air-braked trains were commissioned. After some experimenting with train handling techniques, iron ore trains of 100 wagons became the standard on the Port Elizabeth line.
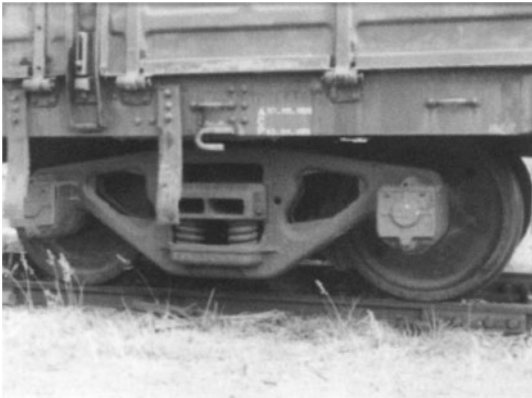
Fig. 1. Sarcast three-piece cast steel bogie

After the introduction of 100 wagon trains on the line to Port Elizabeth which contained many 200 to 250m curves, wheel flange and rail wear was very high, necessitating wheel re-profiling after 30000 to 50000km. The low wheel tread conicities required for an acceptable hunting speed could thus not be retained for long distances due to wheel wear. These worn wheels, however, had an advantage over a new conical wheel tread as far as tread wear is concerned due to a lower contact stress and the fact that the profiled tread causes the contact point to move through a much wider band, thus spreading the wear. These profiles are referred to as *standard wear profiles.*

Recognising the need to maintain the required effective conicity over long distances, a bogie development programme was started in the early 1970s. Initial tests with conventional three-piece bogies soon showed that it was
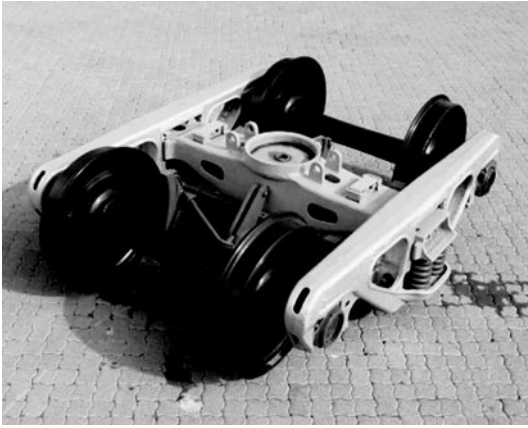
Fig. 2. Spoorbarber bogie

impossible to maintain the *standard wear profile* as wheel tread and flange wear continued at a high rate [2]. Hence, it was concluded that the curving ability of the wheelsets had to be improved. This was achieved by designing a self-steering bogie with a low bending stiffness for adequate curving ability and a high shear stiffness for the required hunting stability.

The new self-steering bogie design retained the three-piece bogie arrangement but introduced rubber shear pads at the journal boxes as well as an additional inter-axle shear stiffness, which is independent of the shear constraint provided by the bogie frame. The first experimental self-steering bogie as shown in Fig. 2 was built in 1971. This bogie used double A-frames (referred to as Double Bissels) to obtain the inter-axle shear stiffness. The coil spring bolster suspension, Barber load sensitive friction dampers and roller bearings of the Spoorbarber bogies were retained.

In the first production bogie the double bissel design was abandoned in favour of sub-frames mounted to the journal boxes and linked diagonally by means of cross-anchors. This design was patented as the Scheffel HS Bogie in 1972. The first Scheffel HS bogies were built in 1975 for the iron ore wagons to be used on the new line from Sishen to Saldanha. A picture of the 20ton axleload HS Mk VII self-steering bogie is shown in Fig. 3. Shortly after its implementation, a 26 ton axleload Mk V bogie was also put into service.

In Fig. 5 and 6 it can be seen that the self-steering bogie basically eliminates wheel flange wear. These measurements were made in 1977 shortly after introducing this innovative design [6]. Furthermore, before the implementation of the self-steering bogie, rails in 300 to 500m curves had to be replaced after about 70 million gross tons when using conventional three-piece bogies at 18.5 tons per axle. Today the original rails on the Richards Bay coal export line are still in use in similar curves after up to 1500 million gross tons, using self-steering bogies running at 26 ton axleload. A significant improve-

ment in the track-friendliness of the three-piece freight wagon bogie was thus achieved by using the self-steering principle combined with an optimisation of the secondary suspension parameters [17,18].

Fig. 3. Double Bissel Bogie

Fig. 4. HS Mk VII Bogie

3   Developments on the Heavy Haul Export Lines

In the following two sub-sections, a detailed account is given with respect to the wheel/rail optimisation strategies followed on the iron ore and coal export lines.

Fig. 5. Wheel profile of a conventional bogie after 20000km

Fig. 6. Wheel profile of an experimental self-steering bogie after 20000km

3.1   Iron Ore Export Line

In April 1976, self-steering bogies at 26ton axleload started to operate between Sishen and Saldanha [7] in trains of up to 202 wagons. The minimum curve radius on the 861km long line is 1000m. After approximately six months of operation it was observed that the wheel wear rate was higher than for the 20ton-axleload bogies. As the wheels were overheating during braking it was initially assumed that the additional tread wear was mainly due to braking. However, a closer investigation showed that the offset of the wheels on the axle was outside the laid down tolerance and the actual rail head profile was different from what was given to the bogie designers.

As early as 1979, an alternating (sinusoidal) side wear on the gauge corner of the rail started to occur [19]. This phenomenon was termed *slinger*, which is the Afrikaans word for stagger. Gouging of the high rail and metal flow on the low rail in curves seemed to be more prominent on maximum gradients where curve radii were less than 2500m. Gouging was also found on straight track but without a particular wavelength. The peak to peak wavelength is typically between 12m and 14m. Figure 7 shows a set of rail profile measurements taken in 1993.

In 1982, 600km of the line had intermittent 2mm side wear, 70km had up to 4mm side wear and 30km had up to 7.5mm side wear. A high rate of wheel wear was also observed on the locomotives and wheels had to be re-profiled after 168000km, then again after 136000km and then again after 80000km. The average locomotive wheel life was thus approximately 400000km while the average wagon wheel life was about 500000km.

Fig. 7. Gauge corner wear as measured in 1993

Following a reduction in maintenance expenditure from 1982 to 1987, and a subsequent increase in traffic volumes from 1988 to 1992 without increased maintenance efforts, the situation reached near crisis levels and the ability of the system to operate at the desired level was under serious threat. Consequently a multidisciplinary approach was initiated to design and implement a combined strategy between the rolling stock and the track engineers [20]. The strategy included:

- Regular locomotive wheel flange and hollow wear measurements.
- Low speed, in motion measurements of hollow wear on the wagon wheels.
- Lateral wheel/rail force measurements to identify skew bogies.
- Rail profiles measurements at selected sites.
- Lateral axle acceleration measurements using an instrumented wagon.

Due to the abovementioned multi-disciplinary approach, the average rate of hollow wheel wear on the wagons improved from 104000km/mm in 1993 to 220000km/mm in 1996 [23]. The main contributing factors were an improved wheel/rail contact band to control sinusoidal gauge corner wear, improved wheel and rail profile designs, pummeling, as well as improved bogie alignment and wheel machining.

However, due to subsequent rail grinding actions, pummeling was again removed with a resulting increase in hollow wheel wear rate to 137000km/mm. Pummeling was then again introduced towards the end of 1999 using asymmetric rail pads to widen the gauge. Presently about 160km is pummelled at +10mm and 40km at +5mm. After the re-introduction of pummeling, the rate of hollow wheel wear has already improved to 167000km/mm as shown in Fig. 8.
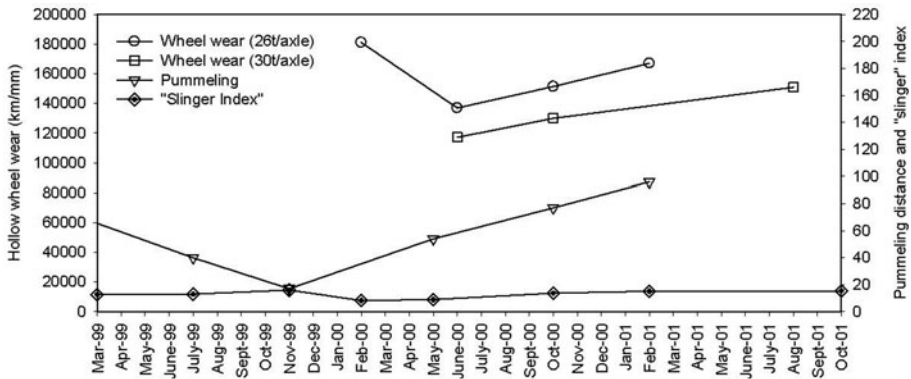
Fig. 8. Influence of pummeling and slinger control interventions

3.2   Coal Export Line

The coal export line was officially opened in April 1976 using 75 18.5ton axleload wagons in a train [21]. In 1979 and 1982 the axleload was increased to 20 and 22ton respectively. In 1981 the train length was increased to 176 wagons. After relaying the track with a heavier and harder S-60 Chrome-Manganese rail, the first 26 ton axleload 164 wagon long trains started to operate in 1986.

Soon after the partial commissioning of the upgraded line, the average wheel life of the leading wheels of the locomotives plummeted from 190000 to 40000km per cut over a period of five months. This resulted in a severe shortage of available locomotives.

The strategy to address this crisis evolved in three phases, the first being one of crisis management, the second one of understanding and the third one of consolidation [22]. A brief summary is given below.

Phase 1 - Crisis Management: During this phase, the rolling stock and the track engineers addressed the problem in a reactive and independent manner by considering the wheel to rail hardness ratio, the prevailing sharp rail gauge corner, the flat rail crown radius, discontinuities at rail joints, a tighter than desired track gauge as well as the absence of track lubrication.

Phase 2 - Corrective Maintenance: The objective during this phase, which stretched from 1990 to 1993, was to achieve conformity between the wheel and the rail. Significant improvements were achieved through gauge control, gauge corner lubrication, rail profile grinding and rail surface defect evaluation.

Phase 3 - Preventive Maintenance: In 1993, maintenance interventions such as rail lubrication, condition monitoring of rolling stock by means of lateral wheel/rail force measurements and a preventive rail grinding strategy were introduced.

In 1993, Tournay presented a paper on experiences gained with the design of wheel and rail profiles for wear control on the coal export line [23]. The following was highlighted:

- The rail gauge corner of the high leg had a very consistent shape.
- The high lateral creep force and the large angle of attack of the leading wheels of the locomotive causes metal flow towards the rail gauge corner.
- The original locomotive wheel/rail contact was mismatched causing severe two-point contact in the initial stages. This resulted in metal flow in the root of the flange and eventually the breaking out of this material.
- Due to high locomotive flange wear, flange lubrication is essential.
- Gauge corner relief was not considered viable as the rate of flange wear on the locomotive wheels indicated that the flange fillet would rapidly wear to the profile of the relieved form of the gauge corner and thus negate this action.
- Minimal flange wear was observed on wagon wheels.
- To limit false flange contact, the hollow wear limit was reduced from 6mm to 4mm.

The above experience led to the decision to use a rail profile very similar to the worn high leg profile. On the low leg the emphasis is on avoiding false flange contact on the field side of the rail. The newly proposed locomotive wheel profile had a high radius differential for curving while the conicity of the wagon wheel profile was lower than the initial profile. The flange of the wagon wheel profile was also made thinner to allow for more lateral motion in order to spread the hollow wear. As much of the surface damage on the high leg was in the form of head checks in the gauge corner, re-profiling of the rail was proposed.

4   Wheel Profile Development

During the development of the self-steering bogie it was realised that with the existing gauge clearance of about 10mm and standard 1/20 conical wheels the rolling radius difference would be limited to about 1mm, which would be inadequate for effective self-steering of the wheelsets. To increase the rolling radius difference it was thus decided to use profiled wheel treads and to increase the gauge clearance to 14mm. This was achieved by a reduction in the back to back distance between the wheels, a reduction in flange thickness and the development of a wheel tread profile which ensures one point contact on curved track. The subsequent chronological development of wheel profiles in South Africa is given in the following paragraphs.

The first profiled wheel profile was the *Standard Wear Wheel Profile*, which developed an adequate radius differential and gravitational stiffness for curving [24]. As the conicity was between 0.25 and 0.42, the operating speeds was limited to 100km/h. Furthermore, high contact stresses occurred

between the wheel and the rail at the maximum lateral deflection of the wheelset. This, together with a high spin component, was considered to be the cause of head checks on the rail.

Hence, it was decided to design a single-radius *Linear Conicity Wheel Profile* with a lower effective conicity and a thinner flange. As a result the hunting stability increased to 160km/h while maintaining a good curving ability. However, high contact stresses still occurred between the wheel and the rail at the maximum lateral deflection of the wheelset. This caused no difficulties under low axle loads, but the profile required adjustment for 18.5 and 26 ton axleload operation.

A parallel development to that of the *Linear Conicity Wheel Profile* was the development of a *Three-Radius Wheel Profile* suitable for 26 ton axleload operation on the UIC 60 rail between Sishen and Saldanha. This profile gave adequate curving ability and hunting stability. High contact stresses however still occurred between the wheel and the rail in the region of the rail crown radius [19]. It was thus decided to develop a wheel profile where contact in the gauge corner of the rail could be avoided.

The resulting profile was the *Drooping Conicity Wheel Profile*, which had a high wheelset conicity of 0.25 when placed centrally on the track and a lower conicity of 0.1 once the wheelset displaced laterally. For dynamic stability the profile behaved as if fitted with 0.12 conicity wheelsets. The high initial conicity enabled an adequate radius differential for curving on the prevailing large curve radii. Hence, contact on the gauge corner of the rail was avoided and the incidence of head checks decreased.

After the implementation of the *Drooping Conicity Wheel Profile* it was found that it was sensitive to variations in track geometry. Hence, a *Modified Drooping Conicity Wheel Profile* was introduced in which a lower sensitivity to track irregularities was achieved by implementing larger radii on the wheel profile in the central region. Hence, the wheelset had to move about 2 to 5mm further laterally to curve. Experience at the end of 1980 showed that this profile best met the set requirements.

Having significantly reduced wheel flange and rail gauge corner wear, the contact stress between the wheel and the rail became an issue [25–30]. The initial *wear regime*, which was characterised by high flange wear rates implied that a wide band of the wheel tread was being utilised. Such wheels are generally worn flat with little hollow wear and have a matching profile between the flange and the rail. Due to this conformal contact lower contact stresses occur between the flange fillet and the gauge corner.

However, due to the improved tracking performance of the wagons, hollow wear started to occur. This resulted in so-called *false-flanges* as shown in Fig. 9. If hollow wear is not limited, *false-flange* contact stresses together with high longitudinal creepages cause longitudinal material flow in the rail resulting in surface shelling and head checks. In sharp curves and under conditions of increased gauge, material flow as shown in Fig. 10 thus occurred

and contact could occur on almost equal wheel radii. To limit *false-flange* contact stresses, hollow wheel wear was limited to 2mm.
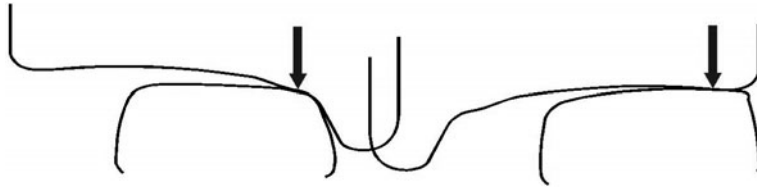
Fig. 9. False flange contact
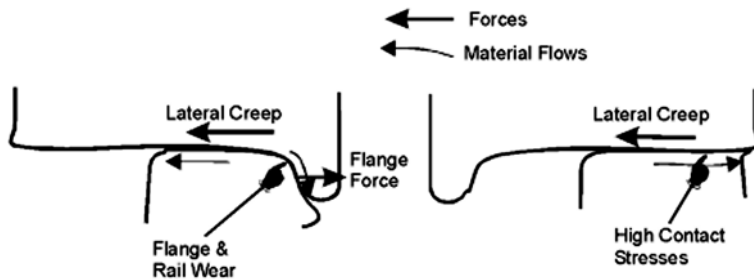
Fig. 10. Forces and material flow on a leading locomotive wheelset in tight curves

Thus, although the *Modified Drooping Conicity Wheel Profile* appreciably reduced the incidence of gauge corner damage, these benefits were lost as hollow wear increased. Hence, a new wheel profile was proposed in 1994. This profile had a convex shape near the flange fillet in order to obtain a sufficient recess of the flange fillet and thus to allow 2mm tread wear before hollow wear would start. This design was abandoned due to high stresses in the area just before the convex portion. To enhance contact to the field side and promote wider contact on the wheel tread, a new *Conformal Wheel Profile* was thus introduced [25].

After implementing the new *Conformal Wheel Profile* together with gauge corner lubrication, rail grinding and lateral wheel/rail force monitoring, the rate of locomotive wheel flange wear reduced in curves sharper than 800m. Furthermore, a preventative rail grinding strategy also evolved to rehabilitate the rail profile.

5   Measurement and Monitoring

To be able to optimise vehicle/track interaction, numerous measurement and monitoring actions were put in place. The following sections give a brief

description of wheel profile, sinusoidal gauge corner wear, pummeling, flat wheel and lateral wheel/rail force measurements, which all form part of the optimisation strategy.

5.1   Wheel Profiles

Wheel profile wear, in particular hollow and flange wear has to be closely monitored to establish any deviations from existing wear rates as well as to be able to determine improved wheel/rail management strategies. These values are used to evaluate the influence of different interventions and to measure the health of the total system.

Wheel wear measurements are conducted using either hand held devices such as 'go no-go' gauges, wheel-life monitoring gauges, the Miniprof electronic measuring system, or by means of track-side mounted plungers or laser beams. Detailed measurements are generally conducted on a selected sample of wheels and occur less frequently.

5.2   Sinusoidal Gauge Corner Wear

To be able to measure the severity of sinusoidal gauge corner wear, the lateral accelerations of the leading wheelset of the leading bogie of a so-called calibration wagon are measured on a quarterly basis on the iron ore export line and annually on the coal export line. Figure 11 contains an example of measured lateral axlebox accelerations of three sequential measurements. The graph in Fig. 12 clearly indicates the reduction in lateral axlebox accelerations due to successful interventions.
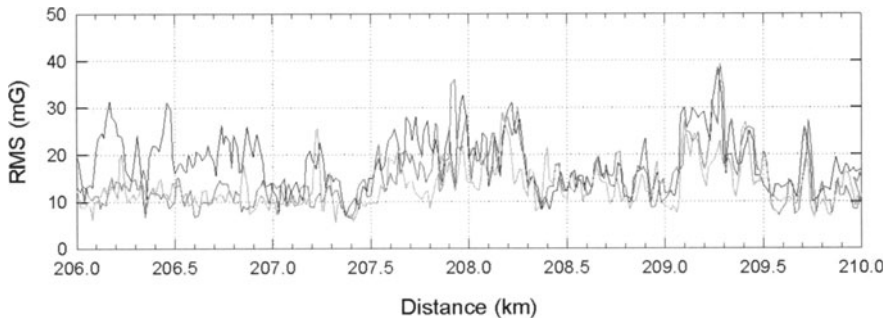
Fig. 11. Lateral acceleration measurements

5.3   Pummeling

To give feedback with respect to the total distance of pummeled track, the distance of pummeled track is measured using a Load Measuring Wheelset.
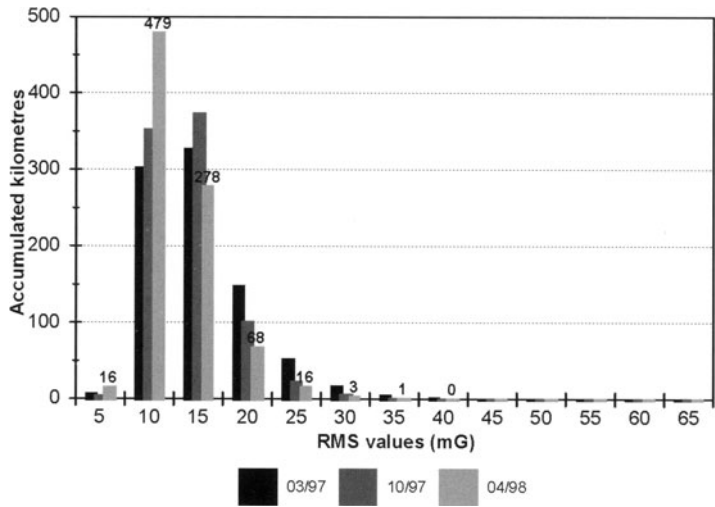
Fig. 12. Reduction in lateral axlebox accelerations due to successful interventions

These wheelsets are able to measure the wheel/rail contact position relative to the taping line of the wheel profile. Figure 13 shows some typical measurements.
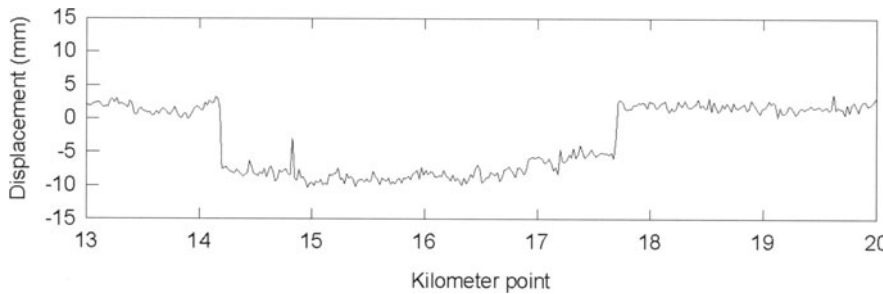
Fig. 13. Measuring pummeled track

5.4   Flat Wheel Detection

Spoornet uses a so-called track-mounted *Weighing In Motion - Wheel Impact Measurement (WIMWIM)* system, as shown in Fig. 14, to identify non-symmetrically loaded wagons as well as to measure wheel impacts due to skidded wheels [31]. The measurement system basically consists of strain gauges mounted on the rail to measure the load of the wheel as it rolls over the crib between two sleepers. All instrumentation can be fitted to the rails without

removal of the rails or interruption to the traffic. With some re-arrangement of sleepers, hundred percent coverage can be obtained for wheel diameters from 730 to 1220mm.

A typical load signal as measured under a flat and a new wheel is shown in Fig. 15. The original signal is then filtered to identify wheel imperfections in relation to undesirable impact loads.

Fig. 14. Track mounted WIMWIM
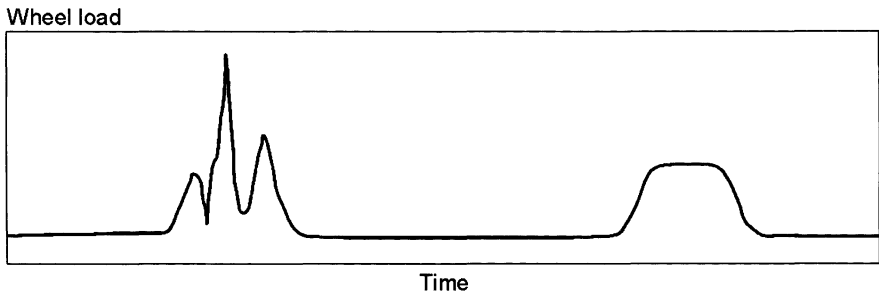
Fig. 15. Signal of original and processed wheel load data

The following limits are used by Spoornet:

- Wheels exceeding a Dynamic Load Factor of 2.1 for 26ton axleload wagons, that is 27.3ton wheel load, must be removed after offloading. At this point, stresses in the sleepers under the rail seat reach their design limit. Skid marks for 26ton axleload operation may thus not be longer than 30mm at the maximum loaded operating speed of 50km/h.

- Due to the quality of rails and sleepers, the recommended alarm limits for the coal and iron ore export lines are 48 and 37ton respectively.
- For vehicles with an axle load between 8.5 and 22ton, wheels exceeding a Dynamic Load Factor of 2.5 should be scheduled for maintenance.

5.5   Lateral Wheel/Rail Forces

The purpose of measuring lateral wheel/rail forces is to identify bogies that are exerting high lateral or bogie couple forces to the track. Such bogies are usually not correctly aligned or have wheels with different diameters within a wheelset. These forces are monitored biannually on the iron ore export line and continuously on the coal export line. An example of normal, gauge spreading and skewing forces is shown in Fig. 16.

   The maximum allowable bogie coupling force for 26ton axleload wagons is 8.5ton. If the bogie couple force reaches this value the train must stop. A bogie couple force of 4.5ton indicates that maintenance must be done. The maximum allowable flanging force is set at 6.9ton.
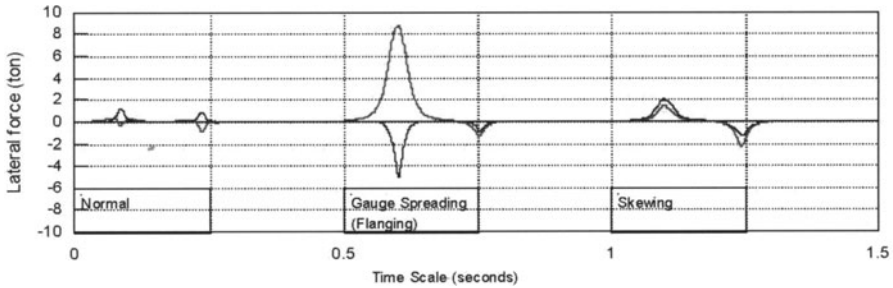
Fig. 16. Lateral forces of a normal, flanging and skew bogie

6   Differential Track Settlement

Another element in vehicle/track interaction optimisation is differential track settlement. This form of track degradation is a function of the spatial variation of the track stiffness as well as the dynamic interaction between the vehicle and the track.

   During research conducted in the 1990s, on-track tests were conducted to simultaneously measure the behaviour of the vehicle and the track [32,33]. A schematic of the instrumented track is shown in Fig. 17. Using the instrumented track, sleeper reaction forces and dynamic track deflection were measured at thirteen consecutive sleepers to determine the dynamic track stiffness as well as changes in sleeper support conditions due to vehicle speed,

axle load and accumulating traffic. Furthermore, Multi Depth Deflection Meters were used to measure the deflection in the various sub-structure layers. In addition to measurements on the track, the dynamic wheel load was also measured using a load measuring wheelset.
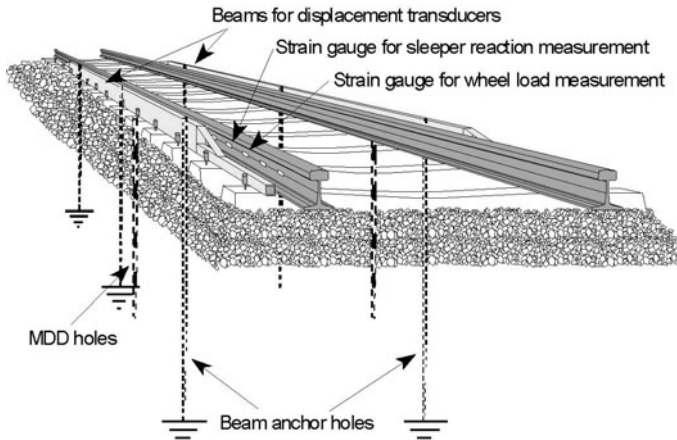
Fig. 17. Test track instrumentation

One of the challenges during the research was to find the relationship between vehicle and track parameters, the dynamic response of the vehicle, and the measured differential track settlement. During the investigation, it was found that the differential settlement of the track was dominated by the spatial variation of the track stiffness.

Using the developed mathematical model, improved fatigue assessment of the track structure is possible, as well as predicting the increase in track roughness as a function of vehicle type, axle load, vehicle speed, track geometry and the structural condition of the track. Realistic dynamic wheel loads are thus available to predict the rate of track component deterioration.

7   Future Projects

A number of projects to further optimise vehicle/track interaction within Spoornet have already been identified. Some of these projects are:

- Establish and document vehicle/track interaction force limits to determine the maximum allowable vertical and lateral wheel/rail forces, or a combination of both, for different track structures.
- Develop a device to measure gauge spreading, gather statistics with respect to gauge spreading stiffness, compare to as-build track and define acceptable limits.

- Determine the mechanism behind sleeper skewing.
- Conduct research to supplement track geometry based measurements with vehicle/track interaction force measurements.
- Predict wheel wear under present and new operating conditions to aid essential proactive research to develop strategies to limit and control wheel wear.
- Develop guidelines for the visual identification and management of wheel/rail interaction problems.

8 Conclusion

From the various aspects of vehicle/track interaction covered in this paper, it is clear that the optimisation of the dynamic interaction between vehicle and track is very involved. The historic overview of the development of the self-steering three-piece bogie, gives the reader a glimpse into some of the elements to be considered when minimising forces or damaging mechanisms acting between the vehicle and the track, as well as between the wheel and the rail. Another important element in this optimisation process is to control wheel and rail profile wear while minimising the contact stress between the wheel and the rail. Relevant optimisation strategies followed on the iron ore and coal export lines are discussed in detail. The description of the wheel profile development as well as measuring and monitoring actions give the reader a better understanding of the interventions to be considered in the implementation and management of this optimisation process.

References

1. Scheffel H. (1972) Hunting stability and steering ability of railway vehicles with particular reference to double-bissel bogie. South African Railways Development of Rolling Stock Research Report DR 15308C, May 1972
2. Scheffel H. (1974) The hunting stability and curving ability of railway vehicles. Rail International, February 1974
3. Scheffel H. (1974) A new design approach for railway vehicle suspension. Rail International, October 1974
4. Scheffel H. (1974) Modified three-piece truck reduces hunting and improves curving. Proceedings of the 11th Annual Railroad Engineering Conference, F.R.A., Railway Locomotives and Cars, Pueblo, October 1974
5. Scheffel H. (1976) Wheelset suspension designed to eliminate the detrimental effects of wheel wear on the hunting stability of railroad vehicles. ASME Symposium on Railroad Equipment Dynamics, Chicago, April 1976
6. Scheffel H. (1976) Self-steering wheelsets will reduce wear and permit higher speeds. Railway Gazette International, December 1976
7. Scheffel H. (1978) Experience gained by South African Railway with the diagonally stabilised (cross-anchor) bogies having self-steering wheelsets. Heavy Haul Railways Conference, Perth, Western Australia, September 1978

8. Scheffel H. (1979) The influence of the suspension on hunting stability of railway vehicles. Rail International, August 1979

9. Scheffel H. (1979) Der Einfluss von Elastizitäten und Kopplungen auf Schlinger-stabilität und Bogenlauf von Schienenfahrzeugen. Leichtbau der Verkehrs-fahrzeuge, 23 Jahrgang, Heft 6, November/December 1979

10. Scheffel H. (1981) The dynamic stability of two railway wheelsets coupled to each other in the lateral plane by elastic and viscous constraints. 7th IAVSD Symposium, Cambridge University, Cambridge, UK, 7-11 September 1981

11. Scheffel H. (1983) The reduction in mechanical wheel rail wear and energy con-sumption as well as total cost benefits when using the Scheffel self-steering truck. International Centre for Transportation Studies (ICTS) Conference, Amafid, Italy, June 1983

12. Scheffel H., Tournay H.M. (1983) An analysis of steady state deviations of wheelsets from the pure rolling position caused by curved track and bogie inac-curacies. 8th IAVSD Symposium, MIT, Cambridge, MA, August 1983

13. Scheffel H., Von Gericke R.E. (1983) The development and design of the Schef-fel self-steering truck and in-service experience gained on the South African Transport Services. Freight Car Conference, Montreal, Canada, October 1983

14. Scheffel H. (1986) Die Konstruktion der Kreuzanker - Drehgestelle der Südafrikanischen Eisenbahnen. ZEV-Glassers Annalen 110, Nr. 6/7

15. Scheffel H. (1996) Unconventional bogie designs - Their practical basis and historical background. Vehicle System Dynamics, Volume 24, Number 6-7, July 1995

16. Scheffel H., Tournay H.M., Fröhling R.D. (1996) The evolution of the three-piece freight car bogie to meet changing demands in heavy haul railroads in South Africa. Proceedings of the Conference on Freight Car Trucks/Bogies organised by the International Heavy Haul Association, Montreal, Canada, June 1996

17. Fröhling R.D. (1998) The influence of friction wedges on the dynamic perfor-mance of three-piece self-steering bogies. 4th International Conference on Rail-way Bogies and Running Gears, Budapest, Hungary, September 1998

18. Fröhling R.D. (1999) The potential for heavier payload on existing track by using track-friendly rail vehicle suspension designs. Proceedings of the World Congress on Railway Research 1999, Tokyo, Japan, 19-23 October 1999

19. Lombard P.C. (1982) Sishen - Saldanha: First six Years. Proceedings of the International Heavy Haul Railway Conference, 1982

20. Kretzmann A.S. (1997) An overview of the multidisciplinary rail/wheel inter-action plan developed for the Sishen-Saldanha line. Proceedings of the Sixth International Heavy Haul Railway Conference, Cape Town, South Africa, April 1997

21. Kuys W.C. (1989) Track structure failure under 26 tonne axle loads on S.A. Transport Services coal line. Proceedings of the 4th International Heavy Haul Railway Conference, Brisbane, September 1989

22. Durham A. (1997) Case study: The coal line wheel and rail interaction strategy. Proceedings of the Sixth International Heavy Haul Railway Conference, Cape Town, South Africa, April 1997

23. Tournay H.M. (1993) Experiences with design of wheel and rail profiles for wear control. Proceedings of the 5th International Heavy Haul Railway Conference, Beijing, China, June 1993

24. Scheffel H., Tournay H.M. (1980) The development of an optimal profile for self-steering trucks under heavy axle load conditions. ASME Rail Transport Division, Winter Annual Meeting, Chicago, November 1980

25. Tournay H.M., Mulder J.M. (1996) The transition from the wear to the stress regime. Wear 191

26. Tournay H.M. (1996) Strategies to control rail and wheel damage: A Spoornet perspective. Proceedings of the Conference on Freight Car Trucks/Bogies organised by the International Heavy Haul Association, Montreal, Canada, June 1996

27. Tournay H.M. (1997) Managing rail wheel interaction. Railways Africa, February 1997

28. Tournay H.M. (1999) Rail/wheel interaction from a track and vehicle design perspective. Proceedings of the International Heavy Haul Associations Conference on Wheel/Rail Interaction, Moscow, Russia, June 1999

29. International Heavy Haul Association. (2001) Guidelines to best practices for heavy haul railway operations: Wheel and rail interface issues. First Edition, May 2001

30. Tournay H.M., Giani J.L. (1995) Rail/wheel interaction: Multidisciplinary practices developed on South Africa. Conference of Railway Engineering, Melbourne, October 1995

31. Tomas M. (1995) Tests to determine maximum allowable size of wheel flat on rail trucks. Spoornet Internal Report, Report No: 24.913/08/95, August 1995

32. Fröhling R.D., Tomas M., Ebersöhn W. (1997) Low frequency dynamic vehicle/track interaction: Instrumentation and measurement. Proceedings of the Sixth International Heavy Haul Conference, Cape Town, South Africa, April 1997

33. Fröhling R.D. (1997) Deterioration of railway track due to dynamic vehicle loading and spatially varying track stiffness. Ph.D. Thesis, University of Pretoria, October 1997

Active Suspension Technology and its Effect upon Vehicle-Track Interaction

Roger Goodall

Loughborough University, Loughborough, Leicestershire, LE11 3TU, UK

Abstract. The paper provides a summary of the principles and concepts of active suspensions, and identifies the particular opportunities for active primary suspensions. It then discusses the major problems and research challenges that exist at present, and finally brings out the key aspects of the technology that will affect the vehicle-track interaction issue.

1 Introduction

For many years railway suspension technology has been principally a mechanical engineering discipline, but now tilting trains in which the suspension response in the roll direction is modified or augmented by the use of control technology are widespread, and these have laid the foundations for a wider use of such "mechatronic" techniques. Aircraft have already made the transition, with fly-by-wire technology now fully accepted, and developments of all-electric aircraft are well advanced. Automobiles are in the process of making the change, with active stability control and other similar systems already in production, and concepts such as "steer-by-wire" being investigated throughout the world. Trains are generally lagging behind the other two transport modes in embracing mechatronics, but there is a range of possibilities: tilting which has already been mentioned, active secondary suspensions and also the general area of active primary suspensions [1]. The paper's focus is upon the third category, which is a general description that can be used to include a number of application opportunities, but is concerned with controlling the wheels and wheelsets in order to improve the basic guidance and stability of rail vehicles.

2 Active Railway Suspension Principles

A conventional mechanical vehicle can be represented as shown by Fig. 1. Inputs are on the left - track inputs, load changes etc, being applied to the mechanical vehicle system. Outputs are on the right: the acceleration levels on the vehicle body, movements of the suspension etc. In a conventional passive system the response is dependent upon parameters such as masses, spring and damper rates, and the suspension geometry. An active suspension brings additional elements to what is already a very complex dynamic system,
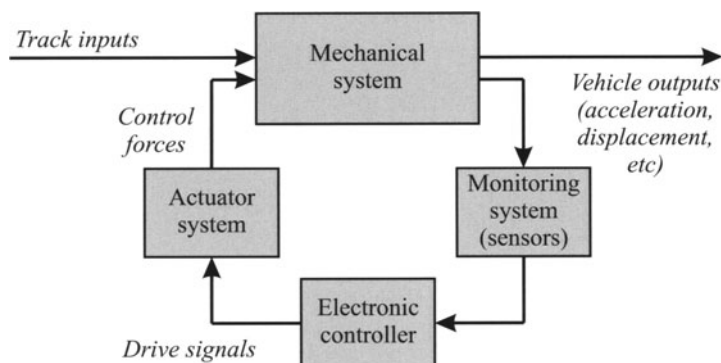
Fig. 1. Generalised active suspension scheme

and its performance will then depend upon what sensors and actuators are used, and upon the hardware and software of the controller. This is therefore feedback control - measure, control and apply corrective action - that's what active suspensions in general are about.

Although active primary suspensions form the main focus for this paper, they are not generally well understood. Therefore it is useful to consider active secondary suspensions (which have been widely studied) as an example to show what can be achieved. The secondary suspension from the bogie to the body is basically there to transmit the low frequency intended movements so the vehicle follows the track, but at the same time to isolate the higher frequency irregularities to provide a good ride quality. An important design problem relates to the level of damping introduced to control the resonance created by the springs. Minimisation of the body resonance effect requires high damping, but because one end of the damper is connected to the bogie it transmits the high frequency effects of the track irregularities. A high damping value therefore makes the ride quality worse by increasing the forces acting on the body at high frequencies, essentially creating a harshness in the suspension response. It creates a difficult trade-off that often includes a spring in series with the damper, and the end result for a typical vertical suspension design is usually around 20% modal damping, quite a low level which means there is still a lot of oscillation in the body response.

If however the dampers are connected to an absolute reference (to give so-called "skyhook damping" - see Fig. 2.) the level of damping can be increased as much as required, and the ride quality doesn't deteriorate [2]. The trade-off between damping the resonance and maintaining a good ride quality has disappeared, although the quasi-static deflections when following the low frequency intended inputs increase so some care is still needed [3]. However application of this concept provides an important improvement in suspension performance.
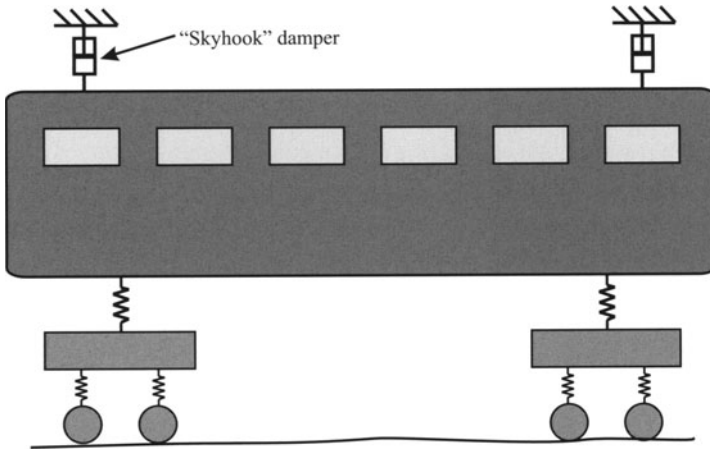
Fig. 2. Skyhook damping concept
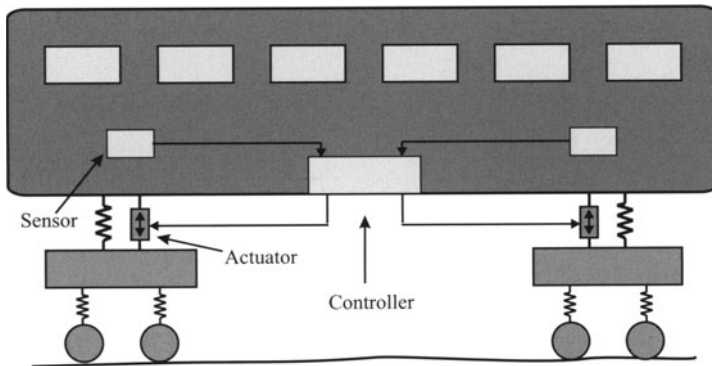
Fig. 3. Active secondary suspension controller

Active control makes it possible to do more than just change the form of the damping. For a normal passive suspension, the spring rates are chosen to give a suitable response in the vertical suspension mode, but having done this the response in the pitch mode is largely determined by the geometry. For a typical suspension geometry the frequency in pitching is usually somewhat higher than in the vertical or bouncing direction, whereas it can easily be shown that there is a definite advantage in having a lower frequency in pitch. It is very difficult to do anything about this for a normal passive suspension, but with an active suspension it is straightforward: the signals from the two ends of the vehicle are brought together, separated into bouncing and pitching motions, and then processed separately to provide control signals for the actuators - see Fig. 3. This way the vehicle modes can be independently controlled, in particular to make the pitching response significantly softer, and the combination of skyhook damping with this "modal control" approach can

be used to halve the acceleration levels on the vehicle body without increasing the maximum suspension movement [4]. The key point of this example is to show the way that active control liberates the process of suspension design. As soon as electronically-controlled suspensions are used it is possible to do things that are not possible with purely mechanical systems - what can be achieved is limited only by what is measured and how many actuators are fitted. An actuator can be made to look like a skyhook damper, even though it is fitted in the place of a normal passive damper, and the "modal control approach" provides an extra design dimension not available in conventional suspension design.

3   The Basics of Wheelset Control

The dynamic properties of the railway wheelset are well understood. The axle connecting the wheels and the coning or profiling of the treads combine to give a natural mechanical steering action. The outward movement of the wheelset when a curve is encountered creates a yawing action such that the curve is followed. The difficulty is that the solid-axle wheelset is also kinematically unstable, and stabilisation is conventionally achieved through introducing a stiff yaw connection, but this also interferes with the curving action. There is therefore a well-known trade-off between stability and curving. There have been many studies into mechanical solutions by which the trade-off can be reduced - yaw relaxation dampers, cross-bracing, self-steering bogies, etc [5]. However the contradiction between achieving high-speed, stable running and effective curving still exists, and in practice results in unnecessary wear of the wheels and rails, and also unbalanced lateral forces between the two axles in a bogie, an effect that may be particularly important on fast large-radius curves. The aim of applying active control to the wheelset is to overcome the contradiction.

3.1   Ideal Curving

It is important first to understand the "ideal" curving position for a railway wheelset. Whether intuitively-derived "classical" or mathematically-formulated "modern" control strategies are used, this understanding is important to guide the process of design.

The starting point is to provide an angle which gives radial alignment with the curve, occurring in opposite directions for the two wheelsets - this is shown on Fig. 4 as α_1. The appropriate angle can be calculated from the geometry using the longitudinal spacing of the wheelsets and the curve radius, ie.

$$\alpha_1 = \ell/R \quad (\ell \ll R) \tag{1}$$
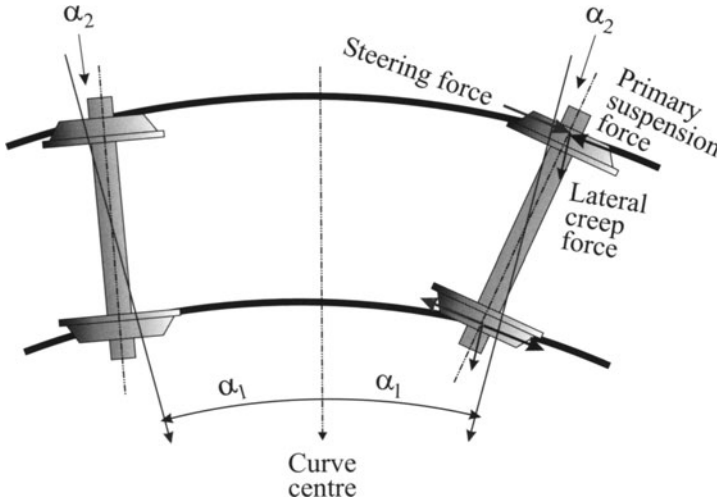
Fig. 4. Ideal curving

where ℓ = semi-longitudinal spacing of the wheelsets, R = curve radius. At balancing speed, ie when there is no cant deficiency, these angles will give ideal curving, but under normal operating conditions it is also necessary to provide the appropriate cant deficiency force, which requires an angle-of-attack, shown as α_2 on the diagram . The angle is dependent upon the lateral creep coefficient, the mass and the cant deficiency acceleration at rail level, i.e. to a first approximation independent of both curve radius and vehicle speed.

$$\alpha_2 = \frac{W \, a_{cd}}{f_{22} \, g} \tag{2}$$

where f_{22} = lateral creep coefficient, W = axle load $[N]$, a_{cd} = cant deficiency acceleration, g = acceleration due to gravity. Whereas α_1 is applied to the two wheelsets in different directions, the angle-of-attack α_2 must be applied in the same direction such that the cant deficiency forces are provided by lateral creep forces on the wheel tread. There also has to be a lateral wheel-rail deflection to give the correct rolling radius difference such that there are no longitudinal creep forces, and hence the yaw torque applied by the creep forces is zero. Accordingly the total steady-state yaw torque applied from the bogie to the wheelsets must be zero. Two consequences arise under this condition: firstly the total wear in the contact patch will be minimised, but also the lateral forces at the wheelsets will be equalised, an important attribute particularly on high-speed curves.

Fig. 5 shows typical wheelset steering angles as a function of curve radius, the middle curve being the purely radial steering case, in other words α_1. The angle-of-attack α_2 is the difference between this and the upper and
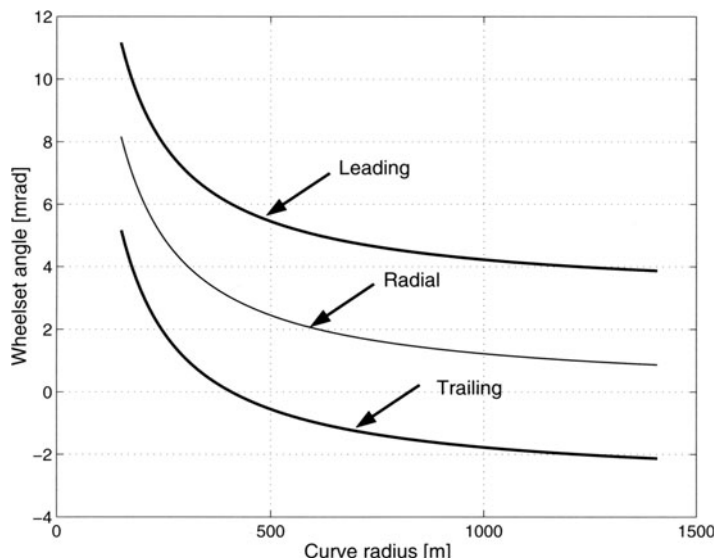
Fig. 5. Wheelset steering angles v. curve radius

lower curves which are for the leading and trailing wheelset angles respectively. The graphs are for a cant deficiency of $2ms^{-2}$, a relatively high value but typical for a tilting bogie. What is important is that the angles are dominated at high speed on large radius curves by the angle-of-attack component, which means that, although radial steering is effective at low speeds where α_1 dominates, it is not effective for equalising the lateral forces on high-speed curves. In fact, to achieve ideal curving, in many cases it is necessary to turn the trailing wheelset completely the other way compared with the radial steering condition - this can be seen from the lower graph, which is negative for radii above 400m.

The discussion so far has been concerned with normal solid-axle wheelsets, but exactly the same steering angles apply for independently-rotating wheels (IRWs). The main difference is that the lateral wheel-rail deflection is no longer constrained to the rolling line - by allowing relative rotation between the two wheels in each wheelset it is possible to satisfy the ideal curving condition anywhere on the wheel tread. One might for example choose to control it so that the wheels are always in the central position, thereby maximising the distance to the flange. However the limiting case is interesting, where the wheelset has moved sideways to run in the root of the flange at exactly the right place for the contact angle to provide the curving forces. No angle-of-attack is then needed, and if the radial steering angles α_1 are applied to wheelsets with IRWs, on a steady curve they will move sideways to exactly the right position, conditional of course on this being a stable operating condition for the wheelset.

3.2   Control Strategies
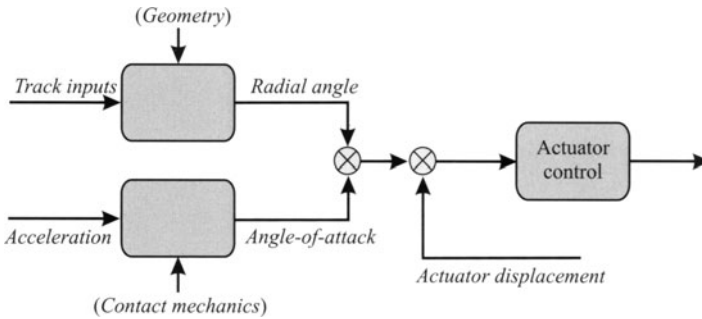
Fig. 6. Feedforward steering controller

The insights expounded in the previous sub-section lead to two different control approaches. The first is principally a feedforward strategy, in which the required steering angle is calculated from an appropriate set of sensors (and other data), and then applied as a command to a steering angle controller - the overall scheme is shown in Fig. 6. This includes feedback around the steering actuator in order to achieve the correct wheelset angle, but essentially this remains a feed forward strategy, and errors in the steering angle command will directly affect the steering accuracy. Such errors will arise from measurement inaccuracies, from not being able to measure everything that is required, or from inaccuracies in the equations used to calculate the steering angle command.
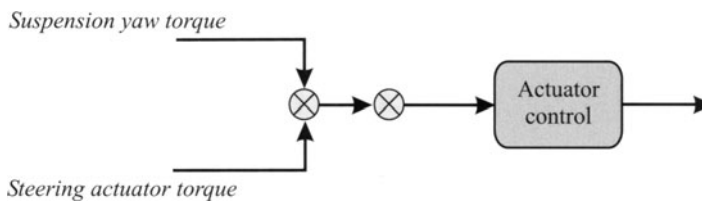
Fig. 7. Feedback steering controller

The alternative is a feedback strategy, derived from the observation that under ideal curving conditions the longitudinal creep forces are zero. The most fundamental approach is to measure and/or estimate the total torque being applied to the wheelset - in Fig. 7 this is shown as the difference between the torques from the primary suspension and the actuator system. Then feedback control can be used to reduce this total torque to zero at low frequencies, typically using integral action. Although this is not directly dictating the wheelset angle, it is setting-up the necessary conditions for ideal

curving, essentially to reduce the longitudinal creep forces to zero. If the wheelset or bogie is stable, this kind of feedback strategy automatically creates the ideal curving condition. This effect was observed a few years ago when the author studied a concept that is called "active yaw relaxation" [6]. The scheme in Fig. 8 shows that the longitudinal link from the axle box includes an actuator in series with the spring; electro-mechanical actuators were used so that at high frequencies there was no movement, and stability was provided by an appropriate choice of series spring. At low frequencies however the actuator was controlled just as described, that is to bring the total force being applied to each axle box to zero.
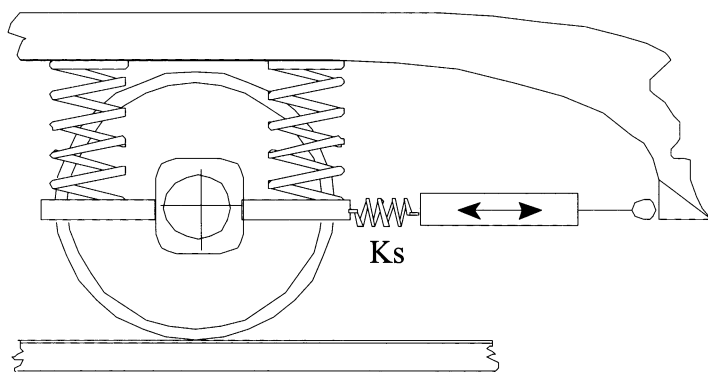
Fig. 8. Active "yaw relaxation" scheme (side view of the arrangement)

Of course this kind of arrangement can be extended so that the actuators provide stability as well as steering, and one stability control approach arises from analysis of a basic wheelset [7]. Although conventional passive dampers are not able to stabilise a bogie, it can easily be shown that active damping is effective. Either a lateral force on the wheelset proportional to yaw velocity, or a yaw torque proportional to lateral velocity, can be used to give stability, both of which can be straightforwardly achieved through active control. Fig. 9 shows how this might be implemented - the signals from lateral accelerometers on the wheelsets are integrated to give velocity, and then a suitable active damping value applied to give a dynamically-varying actuator torque, applied to the wheelsets by means of an appropriate actuator arrangement.

These paragraphs have provided an introduction to some of the possibilities for active steering and stability control, although it must be appreciated that there are a number of theoretical and practical considerations not mentioned here. For example it is necessary to ensure that there is minimal action at low frequencies otherwise the steady-state curving will be upset; similarly the steering action must be slow enough not to affect stability, but both of these requirements are achievable.

Sensors
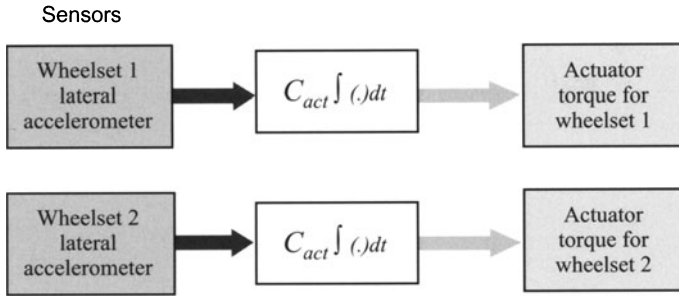
Fig. 9. Active damping controller

3.3   Solid-Axle vs. IRW Wheelsets

Wheelsets with IRWs were mentioned previously, and it is interesting to observe that, once the use of active steering/stability control has been accepted, it makes sense in many ways to use this arrangement rather than the conventional solid-axle wheelset. Of course the natural mechanical steering action has disappeared, but the control can provide this. With IRWs the control action is no longer "fighting" the kinematic motion, and it proves to be much easier to control, and requires lower actuator forces. With a solid-axle wheelset the maximum rolling radius difference still limits the minimum curve radius, even with active control, whereas this is not an issue with IRWs. In fact, purely cylindrical (ie unprofiled) wheels could be used if required, although studies have suggested that sensing of the relative speed between the two wheels is extremely useful for control purposes [8]. Disadvantages are that traction and/or braking are more difficult when the wheels can rotate independently, and also the failure modes may need more careful consideration.

4   Performance of Active-Steered Rail Vehicles

At present, although there are one or two practical implementations, most work is restricted to simulation studies and these give an indication of the kinds of improvements that are possible.

Fig. 10 shows a comparison of wear figures between a passive vehicle and various active control approaches. The detail of the various strategies listed across the bottom is not important - the message to take from these results is that any sensible active strategy can fundamentally solve the problem of wear on curves. There are some differences in performance between the strategies which cannot be seen in the figure, but all give very low levels of wear.

Fig. 11 gives some more results, in this case a comparison between the lateral forces for the two wheelsets in a bogie. The dark column is for the leading wheelset, the light column for trailing. This is a comparison between
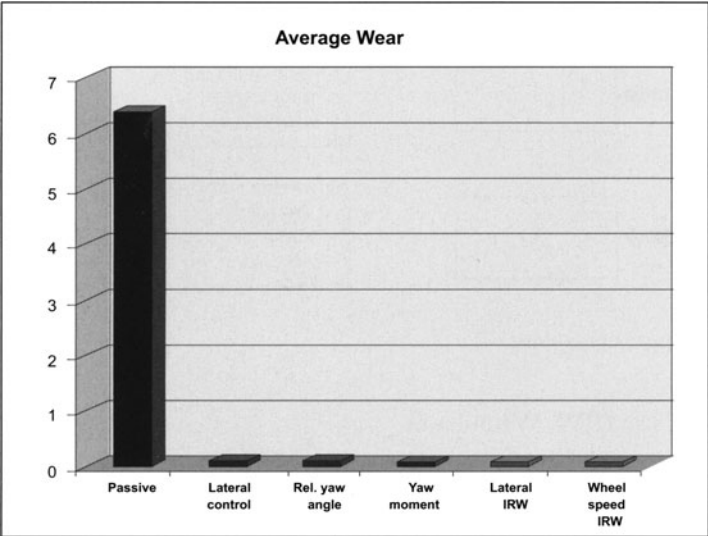
Fig. 10. Improvements in wheel wear through active control
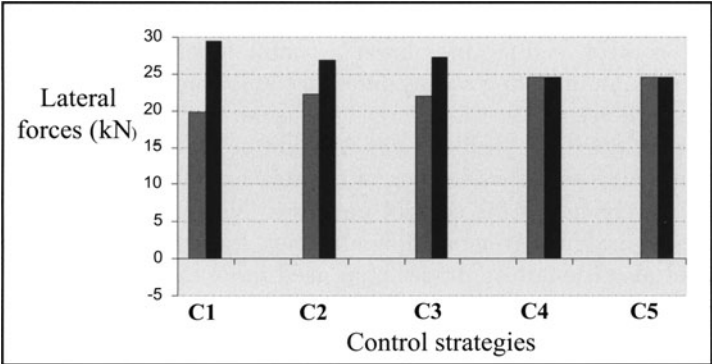
Fig. 11. Equalisation of lateral forces through lateral control

different active steering approaches (C1-C5) and so the passive case is not shown, but it is a rather stiff suspension for high-speed operation and so the lateral force difference for a passive bogie would be significantly larger than for strategy C1. The important point here is that the lateral forces can be completely equalised, as can be seen from the results for C4 and C5.

There is a difference between these two strategies that can only be seen by looking at the longitudinal forces as shown in Fig. 12, which shows the magnitude of the longitudinal force at each wheel - again the dark column is for the leading wheelset, the light one for trailing. For C4 they are zero, but for C5 they are non-zero. In other words, it is possible to devise an active steering strategy that will equalise the lateral forces without minimising the wear, but
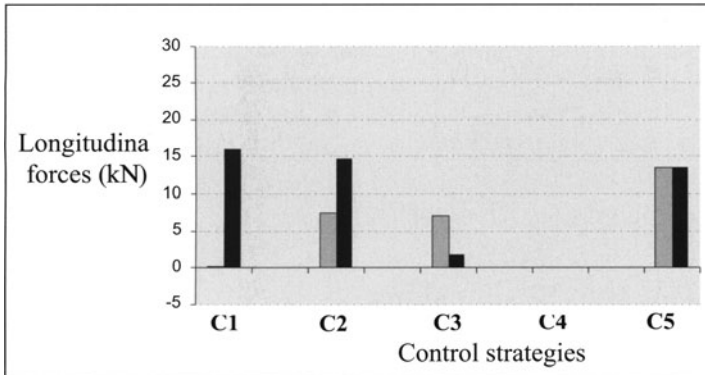
Fig. 12. Minimisation of longitudinal wheel/rail forces through active control

the reverse is not possible - minimising the wear also results in equalising the lateral wheelset forces. Of course there are a number of design issues relating to both steering and stability strategies that have not been addressed here, but these results are sufficient to illustrate the possible performance improvements provided by active control.

5   Research Challenges

Although the control aspects have been simplified to focus upon the principles, a variety of control techniques have been proposed by researchers. However there are a number of key issues that need to be developed. The control engineer always wants lots of sensors when the system is dynamically complex, partly for reasons of basic performance, but more importantly to ensure robustness of the control action as parameters such as conicity and creep coefficients change. By contrast the vehicle systems engineer will want as few sensors as possible, and so the overall control configuration is critical: how many sensors to use, where to fit them, and what types to use are still open questions. There is a similar challenge relating to actuators, not only how many and what technology to use, but also their speed of response. Steering control usually requires rather slow-acting devices, whereas for stability control they must certainly have a much higher frequency range, probably 20Hz or higher. Controller implementation is much less of an issue because of the processing power that is available nowadays. However the controller design process is important: a robust linear time-invariant controller is to be preferred because stability is provable for the vehicle acceptance process, whereas if adaptation is used to accommodate parameter variations, proof of stability is significantly more difficult. Achieving adequate levels of safety and reliability is almost certainly the biggest challenge. Current rolling stock acceptance procedures rely upon a demonstration of high-integrity mechanical components that will not fail under normal conditions. In some cases

this basic principle can be applied with an active system using a mechanical back-up, but generally this will offset a number of the advantages provided by active control, and so new approaches and techniques are needed that will reflect the characteristics of electronic control - highly reliable when designed properly, but with less well-defined failure modes. Inevitably this requires redundancy, fault diagnosis and a structured approach to fault handling and repair - the key attribute being affordability.

6   Impact of Active Suspensions

Although active secondary suspension possibilities have only been briefly discussed, focussing instead upon active steering and stability control, it is useful nevertheless to reflect upon the active suspension possibilities in general in order to provide a perspective upon their relative impacts. The applications have been rather simplistically divided into three: tilting, active ride control and active wheelset control.

Tilting is of course now widespread, but analysis shows that the returns are relatively small: the vehicles are more expensive but give up to 30% higher speed through curves. Generally a lot of track work is required, and then journey time reductions in the region of 12-15% are possible, but without the track improvements this falls to below 10% [9]. Obviously these are approximations because the performance is very route-dependent, but the general message is that tilting requires expenditure on both vehicles an track for relatively small improvements in journey time.

Active secondary suspensions to give improved ride quality are now well understood, but surprisingly little used. It has for example been used on the 500 series Shrinkansen, but only on the end vehicles where there was a problem with ride quality. It is possible that the biggest benefits would come from using the improved suspension performance to provide a good ride quality on power track and thereby save track maintenance costs. At present however the cost base is not well enough understood to be able to make cost-benefit analysis that would confirm or refute this approach.

In contrast, controlling wheels and wheelsets is potentially very large in its benefits compared with these other two options as explained in the next sub-section. The research challenges and engineering hurdles are much bigger than for the first two applications, but perhaps this is inevitable when the impact is much larger.

6.1   Benefits of Active Wheelset Control

There are three possibilities for active wheelset control, all related to vehicle-track interaction, the first two of which follow directly from what has already been said. Firstly, the indications are that with active steering and/or stability control it will be possible to eliminate virtually all "unnatural" wear

of the wheel and rail, in other words the wear that arises from sub-optimum performance of the running gear. It is not possible to do anything about the wear that comes from traction, braking and cant deficiency forces, which is what is meant by the "natural wear". Secondly, the ability with active control to minimise the lateral forces both statically and dynamically on curves is clearly an important attribute. This can be used for the benefit of the vehicle design so that the best possible use is made of the lateral force limit imposed by the track engineer. Alternatively it could be used to achieve lower lateral forces on the track in general, perhaps providing a saving in maintenance cost.
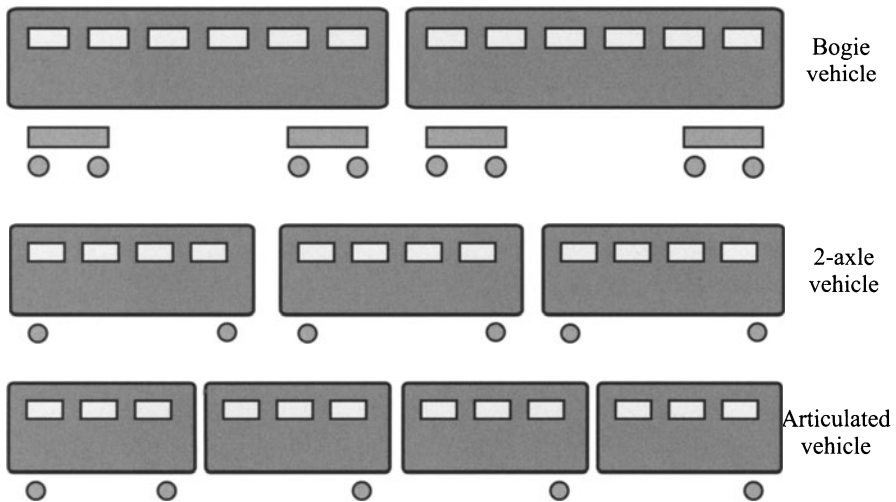
Fig. 13. Simplified vehicle configurations

The third possibility comes from using active control to achieve simpler, lighter, more track-friendly vehicles. Most high-speed vehicles have four axles and two bogies, but two-axle vehicles clearly offer a mechanical simplification and reduction in overall weight due to the loss of the bogie frames. Fig. 13 shows three vehicles instead of two because, although the train weight is reduced, axle load limits will generally mean that halving the number of axles is not possible. There have been examples of two-axle vehicles, but achieving stable running with good curving ability and an appropriate ride quality is difficult with passive suspension, particularly for high-speed operation. However using active suspension technology can make this feasible. In fact this idea can be taken further to consider articulated vehicles: not the TGV solution using articulated bogies, but a more advanced option which probably represents the most simple mechanical configuration possible. Note that the diagram shows a further shortening of the vehicles, again to recognise the axle load limit, but also because underthrow and overthrow on curves is a

constraint. This configuration is similar to the S-Tog trains in Copenhagen, but active control can enable high-speed operation to be achieved, despite the simplicity of the mechanical configuration.

6.2   Vehicle-Track Cost Tradeoff
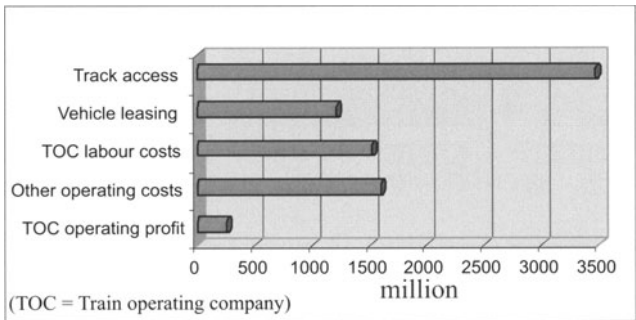
UK rail business: Total =   8 billion p.a.

Fig. 14. Typical annual cost of railway system

The ideas in the previous sub-section are deliberately simplistic, but it is certain that lighter, more advanced vehicles will cause reduced damage to the track. However it is necessary to consider the total cost of railway operation in order to produce a justification. Fig. 14 shows a recent analysis of the total cost of running the UK's railways - subsidies etc are excluded, so this is believed to be the real cost. The high cost of the track is obvious, but it is also necessary to consider the annual cost of the vehicles. Analysis of these two elements in more detail shows that the track cost represents 43% of the total, part of which is for track maintenance amounting to 13%. The annual vehicle cost works out at 15% and the comparison between these two figures is important. The same amount of money is spent on maintaining the track as on the total cost of the vehicles, including their maintenance, and this raises an important question as to whether this tradeoff, which is largely historical, is correct. It seems very likely that there is a good argument for spending more on the vehicles to reduce track maintenance; active suspension technology will add cost to the vehicles, at least in the early days of their development, but it can offer the kinds of improvement in suspension performance that are not possible with other solutions, with consequent reductions in infrastructure costs.

7   Conclusions

This paper has provided an overview of active suspension technology, with an emphasis upon active primary suspensions (ie control of the wheelsets) because these are likely to have the biggest effect upon the railway system as a whole.

Three conclusions emerge:- Firstly, it seems very probable that railways will follow the lead of aircraft and automobiles and move towards increasingly mechatronic solutions. What is less clear is how long this will take to happen, and probably depends upon the determination with which the key research challenges are addressed, in particular the issue of safety and reliability. Secondly, the technology can have a major impact upon the interaction between the vehicle and the track and could result in a re-optimisation of the cost trade-offs between vehicles and infrastructure. Perhaps at some time in the future wheel wear will largely disappear, and rail wear on curves will be a thing of the past? Thirdly, if active suspension technology does become the standard technology for rail vehicles, other possibilities are almost certain to exist, because it is certain that the full potential is not understood at this point in time.

The key question is whether it is possible to engineer safe, reliable and affordable actively-controlled railway vehicles that can fully exploit the technological opportunities in order to release the potentially large cost savings in the infrastructure. Only time will tell!

References

1. Kortum W., Goodall R.M., Hedrick J.K. (1998) Mechatronics in ground transportation - current trends and future possibilities. Annual Reviews in Control **22**, 133–144
2. Karnopp D. (1978) Are active suspensions really necessary? ASME paper, 78–WA/DE/2
3. Goodall R.M. (1993) Performance limits for active secondary railway suspensions. Proc International Conference on Railway Speed-up Technology (JSME) **2**, 81–86, Yokohama
4. Goodall R.M., Williams R.A., Lawton A., Harborough P.R. (1981) Railway vehicle active suspensions in theory and practice. Proc 7th IAVSD Symposium, Cambridge, UK, 301–316
5. Illingworth R., Pollard M.G. (1982) The use of steering suspensions to reduce wheel and rail wear in curves. Proc Instn. Mech. Engrs **196**, 379–385
6. Shen G., Goodall R.M. (1997) Active yaw relaxation for improved bogie performance. PVehicle System Dynamics **28**, 273–289
7. Goodall R.M., Li H. (2000) Solid axle and independently-rotating railway wheelsets - a control engineering assessment of stability. Vehicle System Dynamics **33**, 57–67
8. Mei T.X., Goodall R.M. (2001) Robust control for independently-rotating wheelsets on a railway vehicle using practical sensors. IEEE Control System Technology **9**, 599–607

9. Goodall R.M., Brown S. (2001) Tilt technology still evolving as the cost falls. Railway Gazette International, 521–525

Rolling-Contact-Fatigue and Wear of Rails: Economic and Technical Aspects

Peter Pointner

voestalpine Schienen GmbH Technologie,
Kerpelystrasse 199, A-8700 Leoben

Abstract. High economic and technical requirements need all efforts to increase the knowledge transfer from research and development work to the customer - the railways companies. It is shown that a lot of work has already been done, but there is still a need for further investigation. Research work on Rolling-Contact-Fatigue and its relation to wear and economics is one of the fields of investigation which needs further support by government, industry and railways.

1   Introduction

Rolling-Contact-Fatigue (RCF) -defects are of an increasing importance for railways when planning the maintenance budget. A railway has to spend about 30.000 EUR year by year per km of track. 50% of this money spent are investments and 50% are spent on operation: costs for operational breakdown and maintenance are included here. Especially for repairing and managing RCF-defects 300 Mio EUR are spent per year in the European community.

Any measure is therefore justified to increase track - life; a better understanding of the mechanisms of RCF and wear - taking into account all safety-aspects - will help the railways to save money.
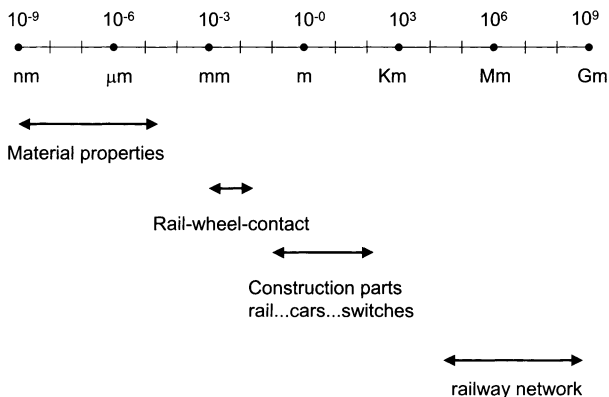
Fig. 1. Length-scale covered in railway systems

2   Our Environment

It is typical of the System Railway that a broad scale of activities are necessary for its evaluation (figure 1). Investigating material properties starts in nm (10^{-9} m) up to μm (10^{-6} m) regions. The size of the wheel - rail - contact area is in the order of 10^{-3} to 10^{-2} m. Calculating and modelling of the wheel-rail system need investigations from some 10^{-2} m to some 10^2 m (size of track components up to dynamic analysis of vehicles in track). Investigations of railway - systems have to be done for up to 10^9 m track length. Therefore the System Railway actually covers a range of 18 orders of magnitude to be investigated. It is obvious that good transfer of knowledge is necessary to cover the whole length scale. Managing the chain from basic research to the customer - the railways - is a very challenging task.

3   Rolling-Contact-Fatigue (RCF)

The rolling of wheels on rails has very specific physical and mechanical characteristics. The elastic modulus of the wheel and rail material is so high that the contact area remains within an order of magnitude of 100 mm^2 inspite of the extensive stresses. These stresses have been investigated thoroughly by a number of simulation programmes, e. g. [1]. The results are very consistent, indicating that these stresses are larger than 1000 N/mm^2 (figure 4). On
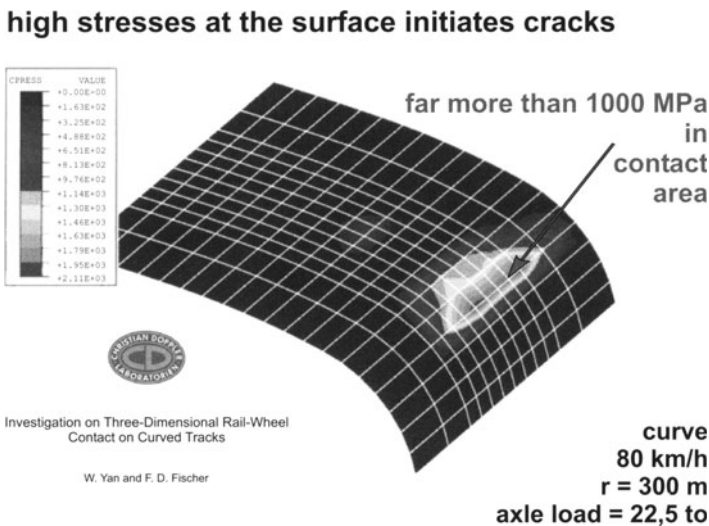
Fig. 2. Stresses in the wheel-rail-contact area

the other hand the material - both wheel and rail - is plastically deformed

and may therefore be subject to damage. The stress conditions may change: having vertical loads and pure rolling of wheels, the interior of the rails can be damaged in the neighbourhood of imperfections (figure 2), whereas - espe-
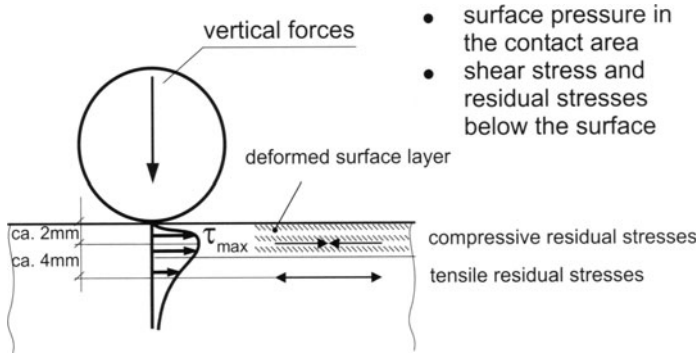
Fig. 3. Load causes material changes

cially when having slip between wheel and rail - RCF-cracks on the surface of the rail may develop (figure 3). A typical example of very severe Head-Checks
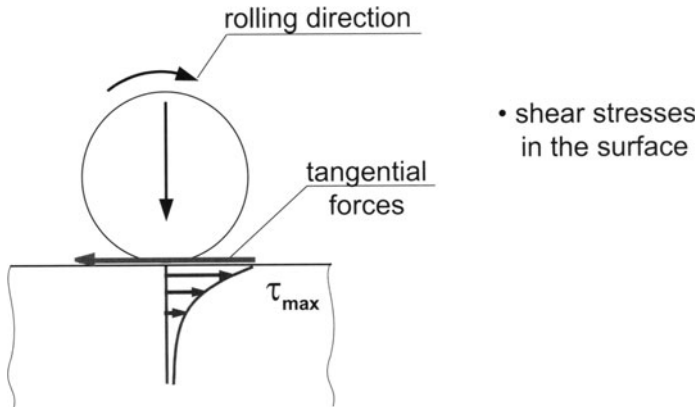
Fig. 4. The forces between wheel and rail - tangential forces

is given in figure 5. In that special case cracks have reached lengths up to some 10 mm: a very detrimental situation.

Material behaviour under cyclic stress is characterized by an accumulation of strain; when exceeding a certain stress level cracks may develop (figure 6). The initiation of cracks under vertical and tangential load is described best by the model of Johnson and Bower [2] in the so called Shakedown - Map (figure 7). Stresses in vertical and horizontal direction cause differerent

No. 919 etched

load ca. 130 MGT's, wear < 2 mm
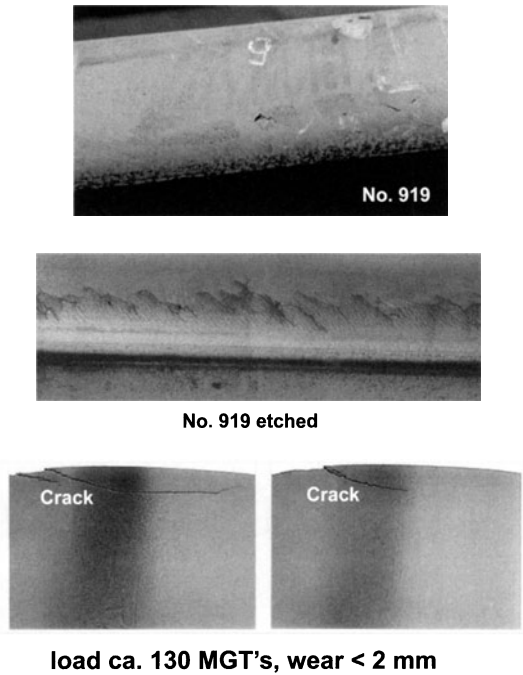
Fig. 5. An example of head-checks developing long cracks
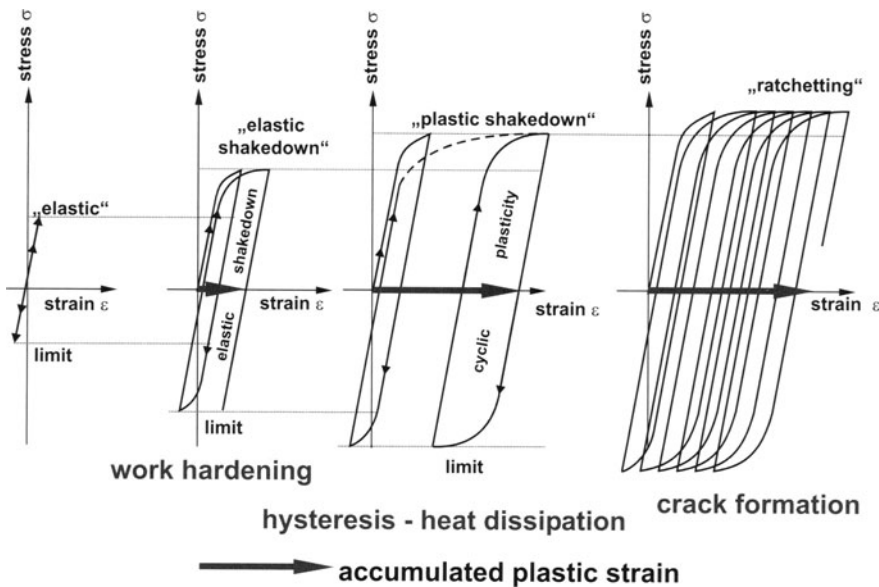
Fig. 6. The extent of damage depends on stress level
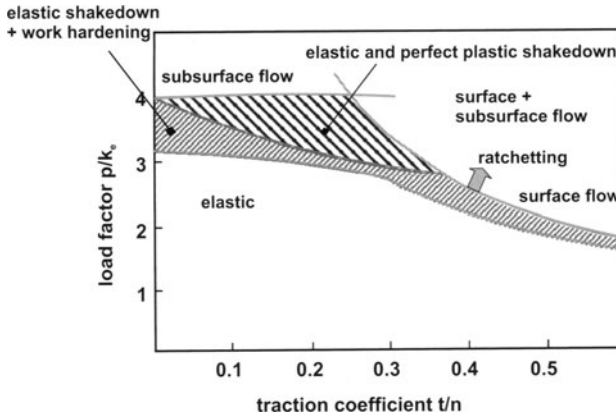
Fig. 7. Shakedown - Map - Description

damage mechanisms: applying vertical stress will lead to subsurface defects - favourized by imperfections in the steel. With horizontal stresses added, which are caused by accelerating or braking actions, the typical RCF-defects like Head-Checks, Spalling and Squats can develop on a smaller stress level.

The development of these cracks is also determined by material properties: the ratio p/k_e on the vertical-axis in the shakedown-map indicates that increasing the shear yield stress k_e decreases the extent of material damage. Some figures for this material value k_e and the stresses calculated for some typical track conditions are given in figure 8.

material			load		
grade	k_e [Mpa]		axle load [to]		p [MPa]
S900A	445		20	tangent	1300
S1100	550		30	tangent	1500
HSH	620		25	curve	2000

k_e: shear yield point [Mpa]   p: pressure in [MPa]
k_e [MPa] $\approx$ 10xHV/6

Fig. 8. Parameters for the Shakedown - Map

This different behaviour of rail-material was observed in a track test: installing and comparing three types of rail - 800, 900 and head-hardened - showed after about one year that Grade 800 had light head-checks, the other two grades had none. In the meantime - after two years of operation - all grades had Head-Checks: inspite of having the highest wear the grade 800

also had the deepest cracks whereas the head-hardenend rails had very fine headchecks.

For managing these cracks it is essential to know something about the history of these RCF-defects. A lot of work has been done recently in the UK; a very useful outcome was to detect a three-stage life of an RCF-crack (figure 9) [3]. Phase 1 is crack development, phase 2 is crack growth at a low
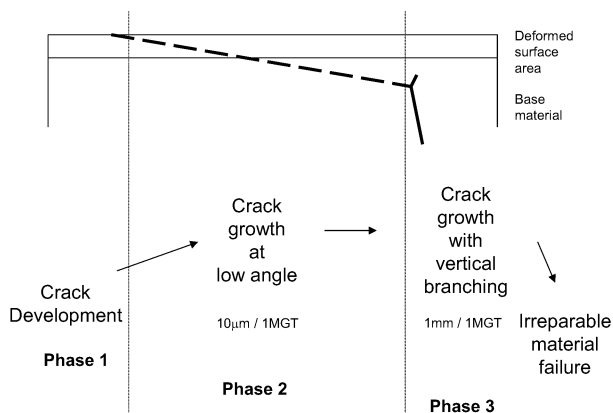
Fig. 9. The local development of a crack

angle into the rail material and phase 3 is the branching of the crack either upward or downward. It is important to realize that the crack growth rate in phase 3 is about 100 times larger than in phase 2. The methods describing the different stages are as follows: for phase 1 the already mentioned shake-down map can be used. For phase 2 more sophisticated models are necessary like FEM analysis or models for low and high cycle fatigue; phase 3 needs pure fracture mechanics applying bending stresses on the rail. Concerning material knowledge there is a lack of data especially for the highly deformed material in the outermost surface area of the rails. The "pots of knowledge" for material data, for loading and stress conditions and for the mathematical and mechanical models themselves are filled to a different degree (figure 10).

This leads to the unfortunate situation that appropriate methods exist for calculating the - already catastrophic - situation in phase 3, but a lack of material data and also insufficient knowledge of the best model for phase 2 make the starting point of phase 3 uncertain.

4   Wear

Wear of rails - and wheels - depends on a lot of parameters, whereof material properties like hardness and carbon-content are only one group. The condition

	Phase 1		Phase 2	Phase 3
Modelling	cyclic / static	Amount of deformation		ok
	Shakedown map		- FEM - Analysis - LCF on plastically deformed material	Fracture mechanics
Data for calculation / modelling				
Data for starting the next phase				We don´t know the starting point of transverse crack growth!
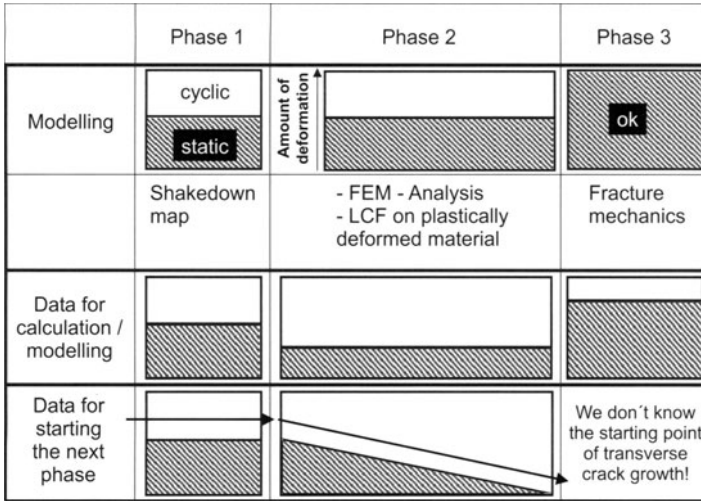
Fig. 10. The "Pots of Knowledge" for calculating RCF-cracks

of the wheel-rail-contact area - dry, wet, lubricated - , the loading conditions - axle load, curves, track irregularities - , and the stress conditions - shape of wheel-rail-contact area, car construction - may play an even larger role in the field of wear.

In our context we only describe the material properties. There was a historical tendency towards harder rail steel; actually we are on the step to introduce the next generation of rail-steel, which can be the harder and tougher bainitic steels. So far, head-hardened rail-steels with superior properties have been used instead of the heat-treated steels according to the CEN-draft (350HT and 350LHT). Examples are the hypereutectoide steel, heaving a carbon-content beyond 0,90% and the low-alloyed steels with a Chromium-content of about 0,50%. These steels provide an improvement in wear resistance of 20% to 100% compared to grade 350HT - heat-treated rails.

One of the advantages of a new generation of rail steels is demonstrated in figure 11: having bainitic steels with higher hardness and tensile strength results in an increased resistance against RCF-defects, providing additionally adapted wear resistance.

5   The Wear - Crack-Length - Diagramm

There are obviously two reasons why rails have to be removed from track: the one is wear and the other are long cracks due to RCF caused by high stresses (figure 12). If the wear is minimised by using highly wear resistant rails and additionally by heavy lubrication, the wear can be so small that cracks can grow deep into the material. If this condition is observed, one of the most
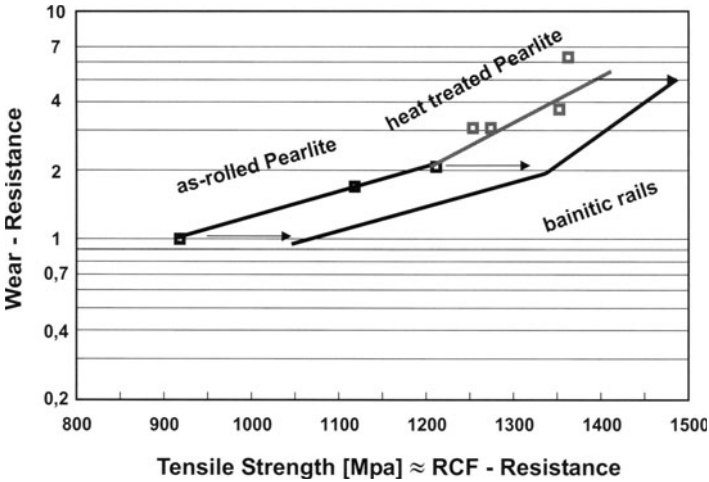
Fig. 11. Increasing wear - and RCF-Resistance

important measures to avoid these cracks is to consider reducing the stresses between wheel and rail. This sounds trivial, but tremendous work has been done how to manage this task: optimizing the wheel-rail shape; best track quality to minimize dynamic effects; roundness of wheels; "trackfriendly" cars;...
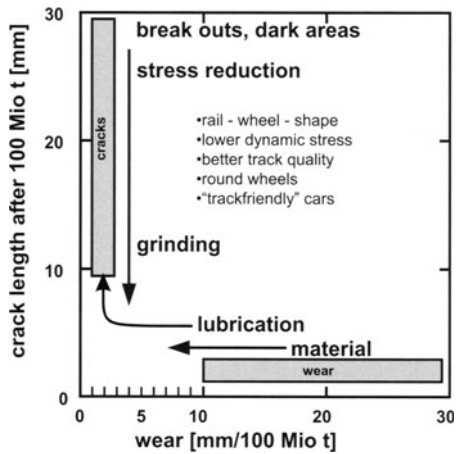
Fig. 12. Part 3: Combining wear and RCF: wear - crack length - diagram

If RCF-defects cannot be avoided by these measures, grinding of rails is necessary to generate artificial wear.

This raises the question for the best method to improve rail life in the railway system : besides all physical and mechanical investigations economic points of view have to be added to research work.

6   Fields of Research and the Link to the System Costs

Basically, three fields of research - according to the length-scale shown above - can be defined for investigating the system railway:

- Basic research describes material properties and contact mechanics;
- Movement analysis describes car- and track-dynamics;
- Operation analysis informs about optimum maintenance strategies how to achieve lowest system costs.

For discussing the interaction between wear and crack growth we will focus here on the third field, which means answering the question: what is the best method of getting maximum track life.

One way of getting the relation between wear and crack growth was shown by British researchers (figure 13) [4]. Rail life is determined by either extensive
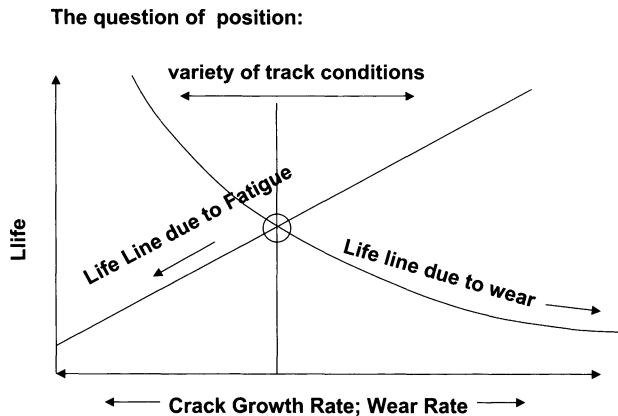
Fig. 13. Part 4: How to manage the system - costs

wear or large crack lengths, both leading to very expensive actions. The optimum point is on the intersection of the two lines. The problem is to find this point. It is not clear, if the intersection remains in the same position during rail-life. Generally the lines are valid for only one particular position in the track; all other conditions can vary considerably.

A similar approach clarifying this problem is shown in figure 14. Considering wear and crack-growth behaviour on a certain position in track, the mechanism will change from higher wear at the begin of rail life to an increasing crack-growth, making grinding necessary. The curve pairs may differ from
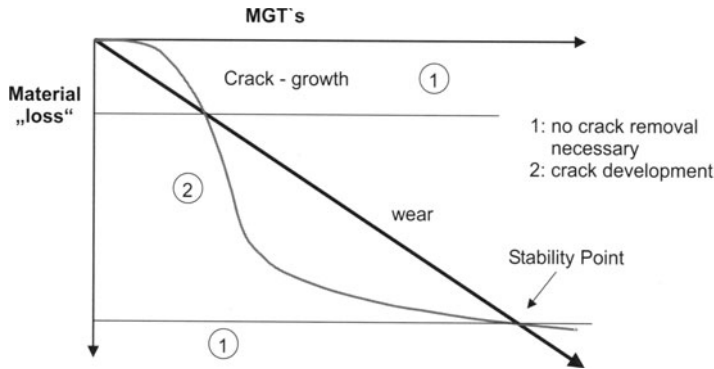
Fig. 14. How to manage the "Self - Wearing - Rail"

position to position, but if the stresses are high enough, cracks will develop and the effect of removing cracks by wear disappears. According to investigations in the UK [3] , the crack growth diminishes after a certain time and a point of stability is reached. It is questionable if this point of stability can be found for all track positions. To manage a system with "self-wearing" rails makes it necessary to know all the stress conditions along the track to find an optimum combination of rail-grades to be installed in in track.
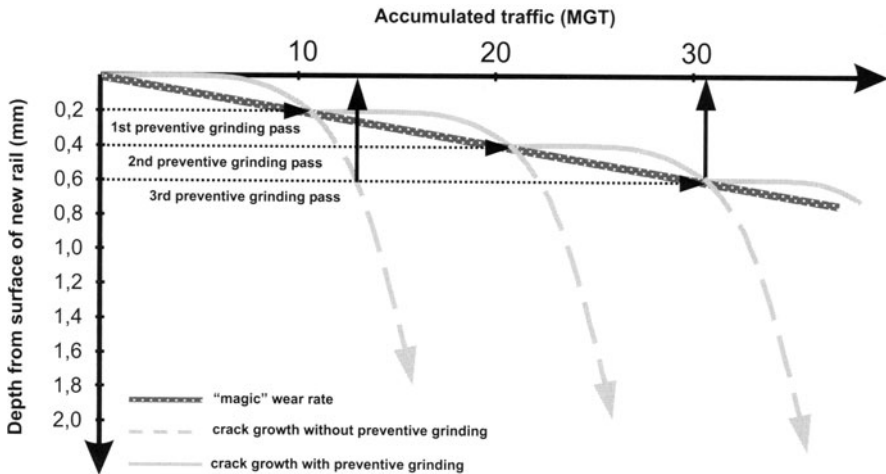
Fig. 15. How to manage maintenance - the "Magic Wear Rate"

A second possibility was introduced by Kalousek and Magel [5]: the principle of the "Magic Wear Rate" (figure 15). As shown above, cracks develop after a - site-dependent - time and grow to a depth making any corrective measures very expensive. The idea of this priciple is to remove the fatigued layers very early and if possible to do it by light grinding. By doing so the

development of cracks starts again on a crack-free rail surface. The so-called "preventive grinding" is repeated periodically. It was shown that it is possible to prolong rail life significantly. This method makes it also possible to use the advantage of head-hardened rails of having a larger resistance against development of RCF-cracks. Understanding this system and managing it properly has proved to offer a safe and cost effective method of maintenance.

7   The Chain from Research to the Customer

As described above, the system railway needs information starting at atomic scale up to the operation of thousands of track kilometers. It should be possible to improve the situation if all basic research work considers the whole system.

A well known problem is sub-optimization. Optimizing for example the costs of the vehicles by simplifying bogie-frame-construction while neglecting about the impacts on the track, finally leads to spending much more money for maintenance. One of the possibilities to solve this problem was shown in the UK. An organisation called WRISA - the Wheel-Rail-Interface-System-Authority - has been founded to direct system optimization [3].

The situation for all basic research work is similar: basic research on material properties, friction management, vehicle- and track- dynamics, development of non-destructive measurement methods.... combined with the investigation of the effect of different materials and maintenance methods on the economics should be done by considering the entire system and the principle that all components influence each other.

Our customer, the railway, is situated at the end of the chain. Research and development work therefore has to provide materials and methods which enable the railway to operate their system at a safe and cost effective level.

The more effort is invested in research work to achieve a positive impact on the system costs, the more it is justified to finance research by government, industry and the railway-operators themselfes. Research at all stages on the length-scale shown above may have a strong multiplying effect; the motto of any research work herewith under discussion should be:

<div align="center">we do research for the benefit of the railway.</div>

References

1. Fischer F., Pointner P. et al. (1997) 3d Finite Elemente Investigation on the Initiation of Head Checks. In: III Int. Conf. on Contact Mechanics, Madrid, 1-3 July 1997
2. Bower A., Johnson K. (1991) Plastic flow an shakedown of the rail surface in repeated wheel-rail contact. Wear 144:1–18
3. Doherty A. (2001) Developing our knowledge. In: National Rail Recovery Programme, Bulletin No 1, May 2001

4. Kapoor, A. (2001) Wear Fatigue Interaction and Maintenance Strategies. In: Why failures occur in the wheel rail system, Venue, Midland Hotel, Derby, 22 May 2001
5. Kalousek J., Magel E. (1997) The "Magic" Wear Rate. Railway Track & Structures 3,50–52

Part II:

Vehicle Dynamics

System Dynamics of Railcars with Radial- and Lateralelastic Wheels

Holger Claus and Werner Schiehlen

Institute B of Mechanics, University of Stuttgart,
Pfaffenwaldring 9, 70550 Stuttgart, Germany

Abstract. This paper presents vibration studies and stability analyses of a railway passenger coach. The wheelset excitations are analyzed to determine which orders of polygonalization generate droning noise in ICE passenger coaches. The strength of conventional wheelsets against vibrations due to polygonalized wheels is investigated. Radialelastic wheels reduce the unsprung mass and isolate the bogie frame and carbody from the medium and high frequency excitation caused by the wheel/rail interaction. A parameter optimization of such wheels leads to considerably reduced carbody vibrations. Stability tests, especially for the so-called hunting motion, are performed for various parameter sets of radial- and lateralelastic wheels. The results show that wheels with increased bending stiffness and improved parameters are feasible, and guarantee the stability of the wheelset motion as well as a noise reduction.

1 Motivation

In 1991 the high speed trains Intercity Express (ICE) were introduced in Germany. During the daily service of the ICE some noise problems appeared. In particular, a medium-frequency structural vibration of the carbody was noticed by passengers after wheels are getting out-of-round, see Pallgen [29]. The so-called ICE-droning is characterized by frequencies in the range of 70 to 100 Hz. First investigations showed that the radius alteration of the running surface of the wheel results in this phenomenon. The formation and accentuation of periodic alterations of the wheel radius is called polygonalization, see Meinders and Meinke [21]. Polygonalized wheels excite the wheelset deterministically and thus lead to the observed vibrations of the carbody. Hence, the vibration transmission between the wheels and the carbody has to be reduced. The most suitable proposal is the substitution of the conventional wheelsets by radialelastic ones where the wheel rim and the wheel disc are connected elastically. Figure 1a illustrates the design of patent [25] realizing rubber-suspended wheels. The noise problems were reduced by this design featuring the so-called third suspension but the strain on the wheel rims increased. As a consequence of strong dynamic loads, a wheel rim broke in the summer of 1998. The train derailed and crashed into a bridge near Eschede, Germany. Immediately, the rubber-suspended wheelsets were replaced by conventional ones. However, two questions arose:

1. Does the axle of a conventional wheelset really resist the dynamic loads caused by polygonalized wheels?
2. Are there other designs for radialelastic wheelsets which provide improved performance and strength while guaranteeing the stability of motion?

The conventional wheelset has to resist static loads from gravitational forces and dynamic loads caused by all kinds of vibrations. The high speed of the train requires the consideration of the rotor dynamics of the flexible wheelset. Moreover, the wheelset may be unbalanced. The proof of the strength of the wheelset requires a computation of dynamic stresses.

The third suspension approach has proved to be suitable for the reduction of the droning noise. Nevertheless, only designs with increased strength against dynamic loads are qualified (cf. Ref. [2]). Here, the patented design [34] is taken into consideration, see Fig. 1b. The wheel rim is reinforced by a T-shaped cross-section which increases the radial bending stiffness considerably in comparison to that of the simple wheel rim in Fig. 1a. The evaluation of the feasibility requires a thorough investigation of the dynamics of a passenger coach equipped with such wheels.

2   Flexible Multibody Systems

The multibody approach, comprising rigid and flexible bodies and using a minimum set of generalized coordinates, has proved to be successful for simulation of mechanical systems, see Ref. [33]. In a flexible multibody system, the rigid bodies represent the inertia and moment of inertia of the stiff system components, and the structural flexibility is modeled by flexible bodies. The bodies may be connected by massless joints and force elements to each other or to the inertial frame. The forces and torques that are acting on the bodies may be volume forces or surface forces. In addition, rheonomic constraints can be introduced by position controls.
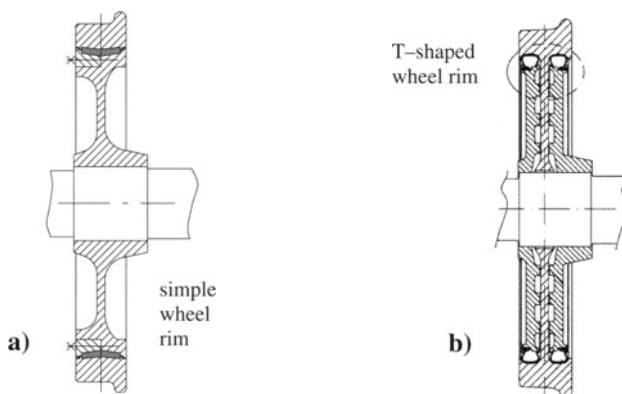
Fig. 1. Wheel rim designs

2.1   Kinematics and Kinetics

In multibody systems a flexible body j may be represented by a reference frame j. This frame describes the large translational and rotational motions which are nonlinear in general. The elastic deformations will be considered relative to this frame. The vector d from the reference frame to a material point in the deformed configuration reads as $d(c,t) = c + u(c,t)$ where c indicates the position of a material point in the undeformed configuration. The displacement field u may be obtained from a finite element discretization as $u(c,t) = N(c)\nu(t)$. The interpolation matrix $N(c)$ comprises the local shape functions, which are each valid within the corresponding finite element only. The vector $\nu(t)$ contains the displacement coordinates of the FE nodes of the flexible structure. The large number of degrees of freedom of the FE nodes may be reduced by modal condensation. The relation between the generalized elastic coordinates $y_e(t)$ and the displacement coordinates $\nu(t)$ of the corresponding elastic deformation can be denoted by

$$\nu(t) = T\, y_e(t)\,. \tag{1}$$

The time-invariant transformation matrix T may be chosen as a reduced modal matrix comprising a set of mode shapes. In this way, the displacement field u is represented by $u(c,t) = \Phi(c)\, y_e(t)$ with the time invariant global shape matrix $\Phi(c) = N(c)\,T$.

   The absolute position of the frame, which is attached to a material point of the flexible body, is given by the vector $r(c,t) = r_j(t) + d(c,t)$. The vector r_j defines the origin of the reference frame j with respect to the inertial frame. The velocity and acceleration of the elastic body can be derived using the relative kinematics. The properties of the flexible bodies may be evaluated with the finite element method (FEM) and stored as time independent volume integrals, the so-called mode shape matrices. The formulation of the Newton and Eulerian equations considers the topology of the system as well as the description of the flexible bodies. Applying the principle of D'Alembert the equations of motion can be written in the form

$$M(y)\ddot{y}(t) + k_c(y,\dot{y}) + k_i(y,\dot{y}) = q(y,\dot{y},t) \tag{2}$$

with the mass matrix M, the vector of generalized Coriolis forces k_c, the vector of stiffness and material damping forces k_i, and the vector of generalized applied forces q. The generalized coordinates of the system are assembled in the vector y comprising the subvector y_r of the rigid body motion and the subvector y_e of elastic coordinates. A detailed description of this approach can be found in Shabana [37]. The simulation of the system can be carried out by standard time integration techniques, see Hairer, Nørsett and Wanner [13].

2.2   Dynamic Stresses

The analysis of the strength of a system component requires the detection of regions of high stress concentration. These regions can be identified in the

stress distributions at a certain sequence of time instants, and representative material points can be selected for further evaluation of stress trajectories. The stress trajectories may be calculated by superposition of normalized stress matrices and the trajectories of elastic coordinates. The extended deformation method described in Seibel [36] and Ref. [6] uses the superposition of a suitable set of eigen-, attachment- and constraint-modes representing also static preloads, see also Craig [9]. If the set of modes is appropriate, not only the static behavior but also the dynamic behavior of the structure is described correctly up to the frequency of the highest eigenmode considered.

The trajectories of all generalized coordinates y_e of the flexible body under consideration are found by MBS simulation. These trajectories provide information about the elastic deformation corresponding to the selected set of mode shapes. Therefore, the multiplication of the transformation matrix T by the trajectories $y_e(t)$ results in the displacement coordinates $\nu(t)$ of all FE nodes of the structure, see (1). The displacement trajectories $\nu(t)$ are used as input for a FE analysis. Then, the stress distribution of the entire FE structure can be evaluated by

$$\sigma(c,t) = H(c)\,B(c)\,N(c)\,\nu(t) \tag{3}$$

where H denotes the Hooke's matrix and B the strain-displacement matrix. Postprocessing by these static analyses results in the corresponding stress trajectories $\sigma(t)$ at each FE node. The advantages of FE postprocessors can be used to visualize stress components, equivalent stresses, or principal stresses of a three-dimensional structure. Thus, even time-variant locations of high stresses can be detected easily.

In the extended deformation method, the stress trajectories $\sigma^k(t)$ at node k are evaluated by superposition of normalized stresses which are scaled by trajectories of the generalized elastic coordinates y_e,

$$\sigma^k(t) = P(c^k)y_e(t) = P^k y_e(t)\,. \tag{4}$$

The stress matrix P^k contains precalculated normalized stresses at the selected node due to the selected set of eigen-, attachment-, and constraint-modes. In accordance with the evaluation of the stress distributions, the results are correct within the limits of the modal condensation. In comparison with the deformation method described in Melzer [22], only a small number of elastic coordinates is required to obtain stresses with high accuracy which is necessary for fatigue limit evaluations (cf. Ref. [3]).

2.3   Stability of Motion

The equations of motion (2) can be transformed to the state space form $\dot{x}(t) = f(x,t)$ with the state vector $x(t) = [\ y^T(t)\ \ \dot{y}^T(t)\ ]^T$ and the vector field f. A solution $x^*(t)$ of this system of differential equations is said to be Lyapunov stable (see Nayfeh and Balachandran [24]) if, for a small number $\epsilon > 0$, there exists a number $\delta = \delta(\epsilon) > 0$ so that any other solution $x(t)$

which fulfills $\| \, x^* - x \, \| < \delta$ for the initial time $t = t_0$ satisfies also the condition $\| \, x^* - x \, \| < \epsilon$ for all $t > t_0$. A solution $x^*(t)$ is asymptotically stable if it's Lyapunov stable and if $\lim_{t \to \infty} \| \, x^* - x \, \| \to 0$. This holds also for the equilibrium solution which is one of the solutions $x(t)$.

The asymptotic stability of a system of linear equations can be checked without the study of the solutions $x(t)$ in time domain but by the investigation of global properties of the equations. The equations of a system with constant coefficients read in state space form as

$$\dot{x}(t) = A \, x(t) + b(t) \tag{5}$$

with the system matrix A. The excitation or disturbance vector b vanishes if the system is not subject to excitation. In this case the system is asymptotically stable if all eigenvalues of the system matrix A have only negative real parts.

Linear systems are not restricted to constant coefficients but may have periodic coefficients. The corresponding equations read in state space as

$$\dot{x}(t) = A(t) \, x(t) + b(t). \tag{6}$$

The stability of such system can be evaluated with the Floquet-Theory if $A(t) = A(t + T)$ with the period T, see e.g. Müller and Schiehlen [23]. The transition matrix Φ can be formulated as $\Phi(t, t_0) = Z(t) \, e^{R[t - t_0]}$ with a periodic matrix $Z(t) = Z(t + T)$ and a constant matrix R. At the initial time t_0 the transition matrix as well as matrix Z become the identity matrix I, $\Phi(t_0, t_0) = Z(t_0) = I$. For $t = (t_0 + T)$ holds $\Phi(T) = \Phi(t_0 + T, t_0) = I \, e^{RT}$. The system is asymptotically stable if the absolute value of each of the eigenvalues of the transition matrix Φ is lower than 1. The transition matrix $\Phi(T)$ can be evaluated numerically through integration of the matrix equation

$$\dot{\Phi}(t, t_0) = A(t) \, \Phi(t, t_0). \tag{7}$$

Local asymptotic stability of nonlinear systems is given if for a solution $x^*(t)$ all solutions $x(t)$ fulfill $\| \, x^* - x \, \| < \delta$ for the initial time t_0 and satisfy $\lim_{t \to \infty} \| \, x^* - x \, \| \to 0$. The set of solutions x is infinite but a finite number of them can be evaluated only. In this way, the asymptotic stability can be tested for a representative subset of solutions through time integration of the equations of motion. This procedure may be very expensive with respect to computation time. However, instability of the system is given if one unstable solution is found.

2.4   Optimization and Parallel Computing

An improved dynamic behavior of a mechanical system can be achieved through optimization of the available design parameters. The optimization requires not only a suitable model representing the dynamics of the system

sufficiently, but also the definition of optimization criteria φ and design parameters $\boldsymbol{p}$ including the admissible parameter range. Additionally, initial conditions for the generalized coordinates of the system must be defined for the evaluation of the criteria. For multibody systems the initial conditions can be written implicitly in vector notation as

$$\boldsymbol{y}_0 : \boldsymbol{\Phi}_0(t_0, \boldsymbol{y}_0) = \boldsymbol{0} \quad \text{and}$$
$$\dot{\boldsymbol{y}}_0 : \dot{\boldsymbol{\Phi}}_0(t_0, \dot{\boldsymbol{y}}_0) = \boldsymbol{0}. \tag{8}$$

After the integration of the equations of motion, the trajectories of the generalized coordinates $\boldsymbol{y}$ are available for the evaluation of optimization criteria. These criteria may be explicit criteria $\varphi^{\mathrm{E}} = \varphi^{\mathrm{E}}(\boldsymbol{p})$ which are independent of the generalized coordinates $\boldsymbol{y}$. Criteria that depend on the dynamic behavior of the system are called functional or integral criteria,

$$\varphi^{\mathrm{I}} = \boldsymbol{G}(t_1, \boldsymbol{y}_1, \dot{\boldsymbol{y}}_1, \ddot{\boldsymbol{y}}_1, \boldsymbol{p}) + \int_{t_0}^{t_1} \boldsymbol{F}(t, \boldsymbol{y}, \dot{\boldsymbol{y}}, \ddot{\boldsymbol{y}}, \boldsymbol{p}) \, dt \tag{9}$$

where the subscript 1 denotes values at the end of the simulation. If a criterion cannot be defined in this way, it has to be defined by the user and, hence, is called user-defined,

$$\varphi^{\mathrm{U}} = \varphi^{\mathrm{U}}(t, \boldsymbol{y}, \dot{\boldsymbol{y}}, \ddot{\boldsymbol{y}}, \boldsymbol{p}). \tag{10}$$

Furthermore, restrictions can be defined as equality constraints or as inequality constraints. During the optimization, the optimization criteria or performance functions, respectively, are to be minimized without violation of the defined restrictions. A multicriteria optimization, see e.g. Eberhard[10], is needed if engineering requirements lead to contradictory performance functions. The optimization algorithm may be deterministic or stochastic. Deterministic optimization algorithms mostly require differentiable systems to compute the next set of design parameters based on the previous set. These algorithms find local minima very efficiently but it depends strongly on the initial design vector which one of the local minima can be found. Stochastic algorithms vary the design parameters stochastically and, therefore, they are not restricted to a local minimum but are able to find even the global one. Nevertheless, stochastic methods imply a large number of criterion evaluations and, thus, are very time consuming.

3   Excitation of Wheelsets

A wheelset is subject to deterministic and stochastic excitations caused by wheel/rail interaction. Stochastic irregularities of the track geometry are described in Popp and Schiehlen[32]. Numerous measurements demonstrate that these irregularities can be understood as stationary stochastic processes

and described by spectral power density functions. The vertical profile ζ is considered as input for the vehicle model presented in this paper. The profile along the distance ξ can be understood as an ergodic Gaussian process, characterized by the mean value $\bar{\zeta}$ and the correlation function R_ζ. The Fourier transformation of the correlation function yields the power spectral density

$$\Phi_\zeta(\Omega) = \frac{1}{\pi} \int\limits_{-\infty}^{+\infty} R_\zeta(\xi)\, e^{-j\Omega\xi} \, d\xi. \tag{11}$$

A spectral representation method, described in Hu and Schiehlen [15], yields efficiently sample functions of the geometrical irregularities from the spectral power density functions. The obtained data represent samples of a realistic track with averaged properties, see Ref. [1].

Deterministic excitation may result from the discrete support of the rail, producing the so-called tie excitation, and/or from the polygonalized running surface of the wheel, respectively. The polygonalization process forms radial irregularities of the running surface which can be described through a series of sinusoidal functions. A wheel has to be reprofiled if the unevenness exceeds a certain amount. By this procedure the wheel radii are quite strongly reduced. Thus, the resulting excitation frequency f_n depends not only on the traveling velocity v and order n of the wheel polygonalization but also, to some extent, on the averaged wheel radius r,

$$f_n(v, n, r) = v \frac{n}{2\pi r} \tag{12}$$

This relation is shown in Fig. 2 for different orders of polygonalization. In addition, the tie excitation is depicted in the diagram. Some relations between frequency and velocity are not visualized by lines but by areas, due to the dependency on the wheel radius. The lower limit of an area corresponds to a
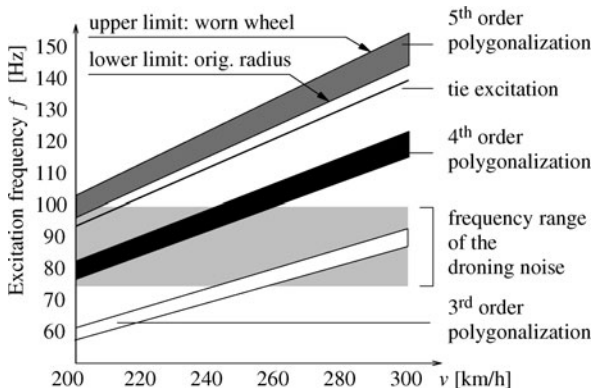
Fig. 2. Possible reasons of vibration by frequency interaction

polygonalized wheel with nearly original radius and the upper limit denotes a polygonalized wheel with minimal radius.

Measurements show that only polygonalized wheels cause the undesired car body vibrations characterized by frequencies in the range of $70 - 100$ Hz. The vibrations occur over a large range of velocity but only at velocities higher than 200 km/h. The frequency range of the undesired car body vibration, the so-called droning noise, is also marked in Fig. 2. Considering linear systems, the excitation frequency has to be within this range too. Exclusively the areas of 3rd and 4th order polygonalization meet this requirement. In accordance, dominant polygonalizations of 3rd and 4th order have been found through measurements made by the German railroad company Deutsche Bahn, cf. Zacher[39]. The polygonalizations may be approximated by sinusoidal functions with a maximum amplitude of 0.3 mm, see Schubert[35]. The tie excitation, which occurs independently of the wheel polygonalization, fits neither the marked frequency range nor any polygonalization of the orders considered. Hence, this type of excitation cannot lead to the above-mentioned droning noise.

4   Strength of Conventional Wheelsets

A model of the ICE passenger coach consists of the carbody, bogie frame, wheelset and rail. Since the vertical dynamics are of major interest here, a plane model with mainly vertical degrees of freedom is considered.
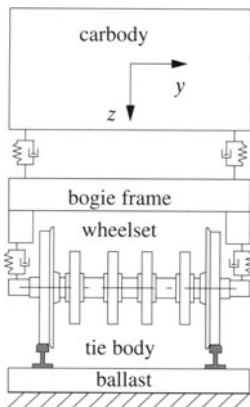
Fig. 3. Vertical model

In this model, the flexible wheelset is forced to run on rigid rails without slip and without yaw motion. Moreover, the bogie frame is prevented from performing a pitch motion. Therefore, the model can be reduced to one wheelset, a half bogie frame and a quarter of the carbody since the excitation of both wheelsets is assumed to be identical. Figure 3 depicts the resulting model. The primary and secondary suspension is realized by linear springs and dampers. It is assumed that the carbody and the bogie frame each have one vertical degree of freedom, and that the bogie frame can also perform a roll motion. The wheelset is in permanent contact with the rigid rails and the model moves longitudinally along the track with constant speed v.

An FE model of the unconstrained wheelset is described by Meinders[20] and shown in Fig. 4. Some modifications are carried out in the FE program ANSYS[26] to prepare this model for further stress evaluations, e.g. reorientation of local coordinate systems. Another delicate problem is the application of the wheel/rail contact forces to the rotating flexible body because the forces have to be applied to FE nodes that are passing the contact
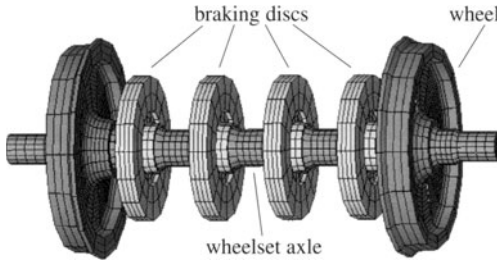
Fig. 4. Finite element model of the wheelset

area. Hence, a contact force would have to be distributed most of the time to more than one FE node and, moreover, these nodes are alternating. A more efficient approach is the transformation of the contact forces to the center of the very stiff wheel rim because this center performs only a translational motion. This approach is based on the fact that the rim is relatively stiff in comparison with the wheel disc whereby the deformation of the wheel rim can be neglected. The normal and tangential forces f_n and f_t at the contact point are transformed into the forces f_n^*, f_t^* and the torque f_m^* acting at the center of the wheel rim, see Fig. 5. In a second step, after considering the rigid rotation motion of the flexible body, the forces are transformed back to the outer link nodes at the very stiff wheel rim by adding four very stiff beam elements, each with one node in the wheel rim center. A similar procedure is carried out for the forces of the primary suspension (cf. Lippold [19]).

The wheelset unbalances are modeled as additional particles attached to FE nodes of the wheels and braking discs as described in Ref. [8]. The original number of about 24,000 degrees of freedom of this FE model is reduced to 1200 through the reduction of Guyan [12]. After the evaluation of the low frequency eigenvectors, the inertia relief modes are computed, corresponding to the applied forces. The elastic properties are transformed to mode shape matrices by the interface program FEMBS [38]. Then the symbolic-numeric equations of motion are generated with the formalism NEWEUL [16].
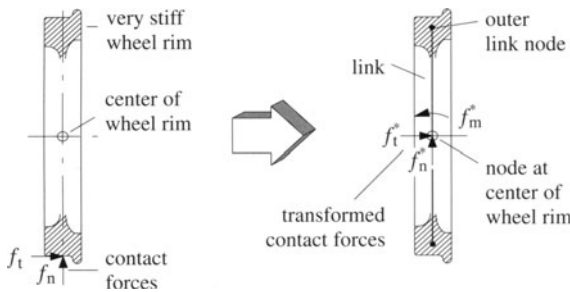
Fig. 5. Transformation of forces and torques
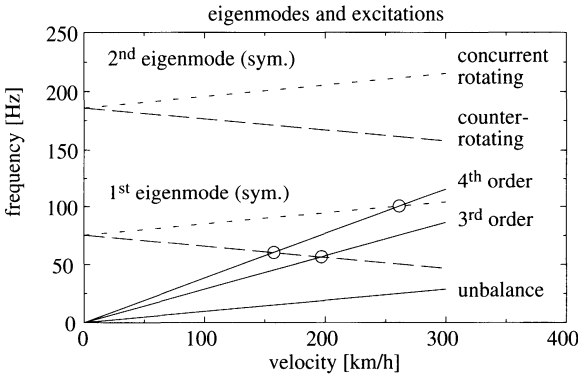
Fig. 6. Frequency graphs of eigenmodes and excitations

4.1   Speeds with Resonance Phenomena

According to Fig. 2, polygonalizations of 3rd and 4th order as well as the unbalances defined above may excite the system considerably. The specification of the speeds at which resonance phenomena occur is difficult because the eigenmodes of a rotating flexible body split into concurrent rotating and counterrotating eigenmodes. The frequency graph of an unbalance is similar to a polygonalization of 1st order. However, unbalances are inner excitations, which excite mainly the concurrent rotating and not the counterrotating eigenmodes. Outer excitations, like those due to polygonalization, affect both types of eigenmodes.

Figure 6 illustrates the splitting of the eigenmodes over the speed range under consideration. Resonance phenomena can be expected in case of deterministic excitations that are close to a system eigenfrequency. Since the excitation depends on the vehicle speed v, the frequency graphs intersect the eigenmode graphs at certain velocities. According to Fig. 6 the intersections of the graphs of the first eigenmode are at velocities of 156.7, 196.1, and 261.3 km/h. The second eigenmode can be neglected since it is not excited here and since the inertia relief modes represent the static deformation sufficiently. The unbalance graph does not intersect the eigenmode graphs at velocities up to 300 km/h. Therefore, the unbalances seem to have only a small influence on the excitation of the wheelset.

4.2   Dynamic Stresses of Wheelset

The simulation program solves the equations of motion and produces trajectories of the generalized coordinates. The deformation of the flexible bodies or data for evaluation of dynamic stresses can be obtained by postprocessing (cf. Peić [31] and Ref. [8]). The applied problem-specific simulation program is based on the NEWSIM library [18] which contains precompiled problem-independent simulation routines.
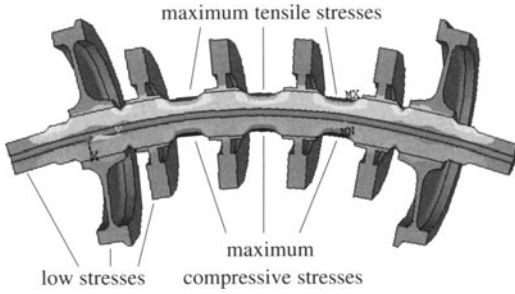
Fig. 7. Stress distribution in wheelset

The first simulation of the vertical model considers a perfectly balanced wheelset running at speed $v = 261.3\,\mathrm{km/h}$ and excited by stochastic track irregularities described in Ref. [1]. The simulation yields maximum tensile stresses of about $65\,\mathrm{N/mm}^2$ that occur between the braking discs. The simulation of the same wheelset, running at the same speed but additionally excited by a polygonalization of 4th order, results in stress trajectories with tensile stress maxima of about $156\,\mathrm{N/mm}^2$. Again, the maximum tensile stress (MX) and the maximum compressive stress (MN) occur between the inner braking discs, see Fig. 7. Further investigations at different speeds show that the highest stresses occur at $261.3\,\mathrm{km/h}$. The simulation of the unbalanced wheelset confirms that the unbalances do not increase the maximum stresses considerably.

For fatigue limit estimations, the material properties of the wheelset are important. The wheelset axle is made of steel. The relatively rough surface results in a permissible alternating bending stress of about $250 - 300\,\mathrm{N/mm}^2$. The maximum stresses of this simple model, which characterize a worst case situation, are still below the fatigue limit and indicate an unlimited lifetime for such a wheelset. More details are found in Pederzolli [30]. An additional investigation is performed at the speed $250\,\mathrm{km/h}$ which is not critical but often used in simulations of the ICE train. The resulting dynamic behavior of a wheelset even with a 4th order polygonalization yields stresses comparable with those of the wheelset without polygonalization. The results demonstrate, however, that resonance phenomena of the flexible wheelset are very sensitive to changes of speed.

5   Feasibility of Radial- and Lateralelastic Wheels

The radial- and lateralelastic wheel consists of very stiff parts like the rim and the disc, and the relatively soft connection between both parts. Therefore, the wheel rims and wheel discs can be considered as rigid bodies. The flexible connection in between is represented by linear, massless springs and dampers, see Fig. 8, with the suspension coefficients $c_{\mathrm{RE,y}}$, $d_{\mathrm{RE,y}}$ in lateral direction and

$c_{RE,z}$, $d_{RE,z}$ in radial direction. Each wheel rim may undergo longitudinal, lateral, vertical, roll, and yaw motions (x_{RE}, y_{RE}, z_{RE}, α_{RE}, γ_{RE}) relative to the wheel disc. It is assumed here that there is no major torsional flexibility. The wheelset axle is modeled as rigid body. The radial- and lateralelastic wheelset is related to the conventional one so that their masses and inertias correspond to each other.
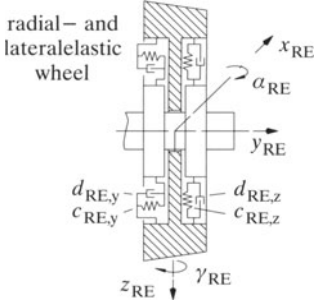
Fig. 8. Model of the wheel

The stiffness and damping properties of the third suspension can be approximated by springs and dampers which are fixed at coordinate systems of the wheel rim and disc, see Fig. 9a. This approach realizes general damping of the rim motion (confer Ref. [5]). The equations of motion can be linearized at the equilibrium position and transformed into state space. Then, the corresponding system matrix A has constant coefficients wherefore the stability of the system is characterized by the eigenvalues of the system matrix.

A more realistic damping model can be achieved with springs and dampers which are attached to material points, see Fig. 9b. This specific damping results in equations of motion with periodic coefficients depending on the period T of the wheel rotation. Linearization at the equilibrium position and transformation into state space yields a time dependent system matrix A with periodic coefficients. The application of Floquet's theory can be used to get information about the stability of the system, see Section 2.3.

The system topology and the system parameters determine the equilibrium position. The example of a rotating wheel, with disc mass m and linear springs and dampers, illustrates the difference between general and specific damping, when exposed to the gravitational acceleration g. The only degrees of freedom are x and z. The equilibrium position of the model with springs and dampers attached to coordinate systems (Fig. 9a) depends on the stiff-
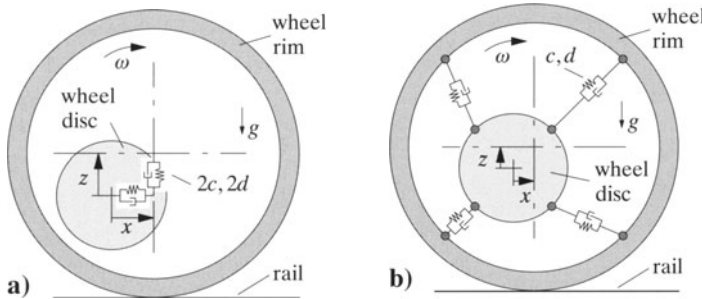
Fig. 9. Modeling of suspension and damping

ness c, the index g indicates general damping,

$$x_{\mathrm{E}}^{\mathrm{g}} = 0 \quad \text{and} \quad z_{\mathrm{E}}^{\mathrm{g}} = mg\,\frac{1}{2\,c}\,. \tag{13}$$

The model with springs and dampers attached to material points (Fig. 9b) possesses an equilibrium position depending not only on the stiffness c but also on the damping d and the angular velocity ω,

$$x_{\mathrm{E}}^{\mathrm{s}} = mg\,\frac{d\omega}{2(c^2 + d^2\omega^2)} \quad \text{and} \quad z_{\mathrm{E}}^{\mathrm{s}} = mg\,\frac{c}{2(c^2 + d^2\omega^2)}\,. \tag{14}$$

The specific damping is denoted by the index s. Equilibrium positions $x_{\mathrm{E}}^{\mathrm{g}}$ and $z_{\mathrm{E}}^{\mathrm{g}}$ as well as $x_{\mathrm{E}}^{\mathrm{s}}$ and $z_{\mathrm{E}}^{\mathrm{s}}$ trend to zero for increased stiffness. The behavior of a rigid wheel is given for $c \to \infty$. Low damping or low velocity leads, for any stiffness coefficients, to similar equilibrium positions of both models.

The derivation of analytic expressions for the equilibrium position of a radial- and lateralelastic wheel is much more difficult due to the dependency on system parameters and velocity. Alternatively, this position can be evaluated numerically. In this way the linearization of the nonlinear equations of motion would have to be performed for each set of parameters and each desired speed. Nevertheless, the integration of the equations of motion is found to be very efficient for stability tests if suitable initial conditions are used.

5.1 Vertical Dynamics

The model of Sec. 4 does not show the vertical vibrations resulting in droning noise; it must therefore be adjusted as described in Ref. [4]. The bogie frame is modeled as a flexible body to include all medium-frequency resonance phenomena. Here, the first three modes of the half bogie frame model as well as the eigendynamics of the secondary suspension are considered. The model comprises either rigid wheelsets or radialelastic wheels with general damping.

The evaluation of the sound pressure level due to the carbody vibration would require an even more detailed model. Therefore, the spectral density Φ of the carbody acceleration $\ddot{z}$ is used as an indication of the magnitude of the droning noise. The spectral density function, which depends on the vibration frequency f, is obtained from the correlation function $R_{\ddot{z}}$,

$$\Phi(f) = \frac{2}{\pi}\int_0^{+\infty} R_{\ddot{z}}(\xi)e^{-j2\pi f\xi}d\xi\,. \tag{15}$$

The simulation results are shown in Fig. 10. The model with rigid wheelsets leads to increased spectral densities mainly in the frequency range from 80 to 100 Hz. These increased amplitudes indicate the carbody vibration which leads to droning noise. The results of the model with radialelastic wheels show that the acceleration of the carbody decreases considerably in the critical frequency range between 70 and 100 Hz, see Hirsch [14].
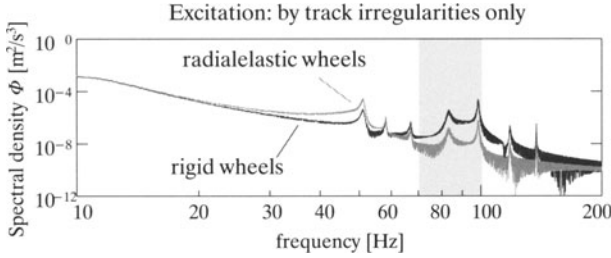
Fig. 10. Spectral density of vertical acceleration

The reduction of the droning noise requires a suitable choice of the suspension coefficients while the wheel/rail forces should at least not be increased. Accordingly, the vector of design variables includes the coefficients of the radial elasticity, the primary and the secondary suspension. The optimization goals have to be formulated as criteria. In this way, the mean square value of the carbody acceleration, restricted in the frequency range of interest, is chosen as a measure of the magnitude of the medium-frequency vibration,

$$\sigma_{R,\ddot{z}}^2 = 2\pi \int_{70\,\text{Hz}}^{100\,\text{Hz}} \Phi(f)df \ . \tag{16}$$

High dynamic forces between wheel and rail increase wear and may even result in derailment. Therefore, the mean square values of these forces are used as second optimization criterion. A decreasing stiffness of the radial elasticity will reduce the carbody accelerations, but it leads to an increasing static deflection of the wheel rim, too. The lower limit of the radial stiffness coefficient $c_{RE,z}$ is set to 10^7 N/m to restrict the static deflection to 5 mm at most. The initial design parameters are taken from the wheelset with a simple wheel rim. It is assumed that the parameters may vary between 0.1 and 10 of these initial values. For a more efficient simulation, the sophisticated model secondary suspension is no longer considered in the optimization.

The parameter optimization is carried out by the program MCDM (Multi Criteria Decision Making[10]) which provides various optimization algorithms, multicriteria methods and routines for the evaluation of the criteria. The program interacts with the software package PVM (Parallel Virtual Machine[11]) which calls the criterion evaluation program on the available computers, transmits the design parameters and receives the criterion values. A detailed description of the implementation is given in Kübler[17].

Figure 11 shows that the optimized parameters reduce the spectral density of the carbody acceleration strongly in comparison to those of the initial design or of the model with rigid wheelset. In the frequency range between 70 and 100 Hz the average of the carbody acceleration is reduced to only one percent of the original value. Moreover, the optimized parameters yield a very low force level on the wheel rim. More details can be found in Ref. [4].
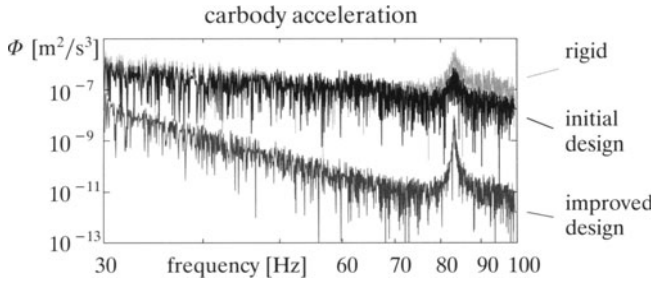
Fig. 11. Comparison of initial and improved design

5.2 Stability of Motion

The spatial model for the stability analysis, see Fig. 12, consists of a half carbody and a rigid bogie equipped either with two rigid or two lateral- and radialelastic wheelsets, respectively. The running wheelsets undergo lateral, vertical, yaw and roll motions. The rim of an elastic wheel has five degrees of freedom according to Fig. 8. The springs and dampers are attached either to coordinate systems or to material points, see Fig. 9. The lateral motion of the wheelsets affects the carbody motion only slightly. Hence, the carbody motion is constraint laterally. The straight track is fixed to the ground. The geometric profiles of the wheel rims are conic with the conicity $\delta_0 = 0.35$ rad, the cross section of the rail head is assumed to be rectangular. These assumptions reduce the contact zone of wheel and rail to a single point. The creepage is derived from the nonlinear kinematics and, thereafter, linearized with respect to the generalized coordinates. The creep forces are evaluated using Kalker's linearized theory. The normal forces acting on the wheelset are represented by very stiff springs (confer Ref. [7]).

   Below the critical speed a conventional wheelset performs a damped hunting motion. Thereby, the wheelset returns to its central position after lateral disturbance. At the critical vehicle speed the hunting motion becomes a steady-state oscillation. Consequently, the hunting motion gets unstable for speeds higher than the critical one. The analysis of the eigendynamics of the model with conventional wheelsets confirms that the critical eigenvalue characterizes the hunting motion. An evaluation with the computer algebra program MATLAB [27] results in a linear critical vehicle speed of approximately 360 km/h. Moreover, the root locus plots show that the hunting motion has the lowest frequency of all periodic eigenmotions of the system, see Oberle [28]. The model can be simplified by exclusion of the lateral and yaw motion of the bogie frame and, in this way, only one wheelset has to be considered. Accordingly, the longitudinal stiffness of the primary suspension is modified in order to achieve a critical speed of 360 km/h for the simplified model, too. The critical speed can be verified easily by simulation if the initial conditions are close to the steady-state solution. In this case, the initial
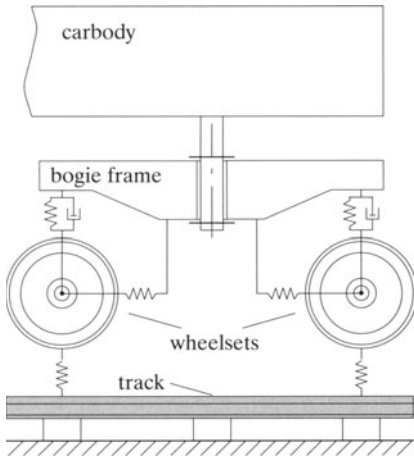
Fig. 12. Spatial model of passenger coach

conditions don't excite motions with higher frequency than the hunting motion. For speeds above the critical one the amplitudes of the hunting motion increase, indicating that the system has no stable equilibrium position any more. The lateral amplitudes of the wheelset relative to its central position decrease for speeds lower than the critical one and, therefore, the system is assumed to be stable. The evaluation of the critical speed for different values of design parameters is automated with the FORTRAN interface of NEWSIM [18].

The above described model with one wheelset is extended to lateral- and radial elastic wheels with general damping of the wheel rims. The parameters for the radial elasticity are set to $c_{RE,z} = 1.0 * 10^7$ N/m and $d_{RE,z} = 5.0 * 10^4$ Ns/m according to those of Sect. 5.1. Thus, the coefficients $c_{RE,y}$ and $d_{RE,y}$ of the lateral wheel rim suspension are the only design parameters. A parameter study with respect to the critical speed and over a large range of stiffness and damping parameters is carried out. Figure 13 shows that the linear critical speed for high stiffness and damping values is close to the critical speed of 360 km/h of the conventional wheelset. The critical speed is increased at lower stiffness and intermediate damping coefficients. There, the increased steady-state inclination of the elastically suspended wheel rims reduce the effective conicity of the wheel profile and worsens the curving behavior of the wheelset. The destabilizing effect of the lateral motion of the wheel rims dominate at lower damping values and decrease the critical speed. Figure 14 shows the variation of the vertical stiffness and damping coefficients $c_{RE,z}$ and $d_{RE,z}$. The lateral coefficients are set to $c_{RE,y} = 1.0 * 10^8$ N/m and $d_{RE,y} = 1.0 * 10^3$ Ns/m. The results illustrate that the optimal parameters for low droning noise are robust with respect to the stability behavior of
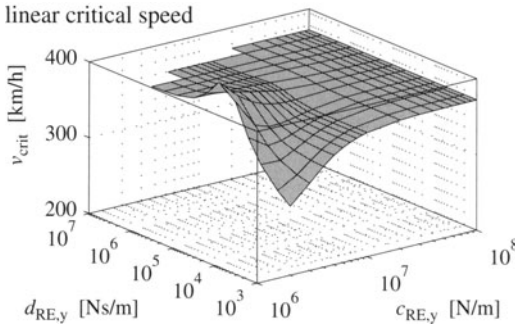
Fig. 13. Variation of lateral elasticity, general damping
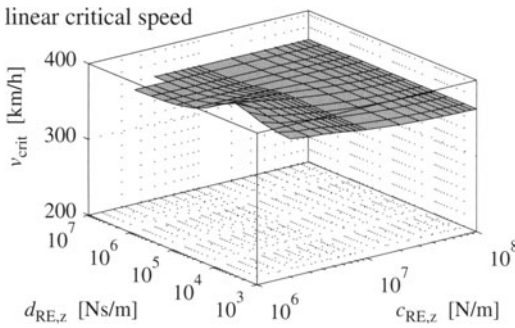
Fig. 14. Variation of vertical elasticity, general damping

the running wheelsets. The critical speed is only slightly increased at lower stiffness and damping values.

Now, the general damping of the lateral- and radialelastic wheelsets is replaced with the specific damping where the springs and dampers are attached to material points. The results of variation of the lateral elasticity, see Fig. 15, are similar to those with general damping but the maximum of the critical speed is lower, and located at smaller damping values. The rotating dampers stabilize the wheel rim motion at these high speeds. Anyway, the curving behavior may be worsened at lower speeds in case of low lateral stiffness and damping values. The variation of the radial elasticity is carried out for $c_{RE,y} = 1.0*10^8$ N/m and $d_{RE,y} = 1.0*10^3$ Ns/m. Figure 16 shows that no equilibrium position can be found at low stiffness and high damping values. In this range an overdamped motion of the wheel rim occurs and increases the eccentricity of the wheel rim relative to the wheel disc.

The investigation of radial- and lateral wheels turns out that the stability behavior is very robust for lateral stiffness $c_{RE,y} \geq 5.0 * 10^7$ N/m. Hence, the lateral stiffness has to be considerably higher than the radial stiffness $c_{RE,z}$ to guarantee stability. The damping coefficient is an uncertain design parameter which varies in a relatively wide range because it is very sensitive
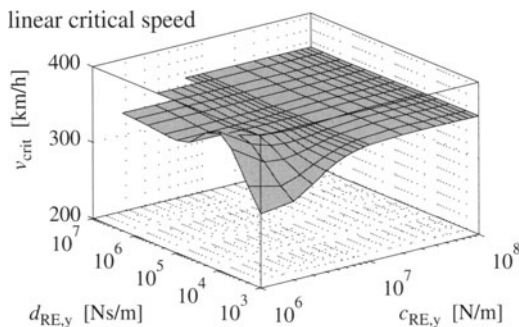
Fig. 15. Variation of lateral elasticity, specific damping
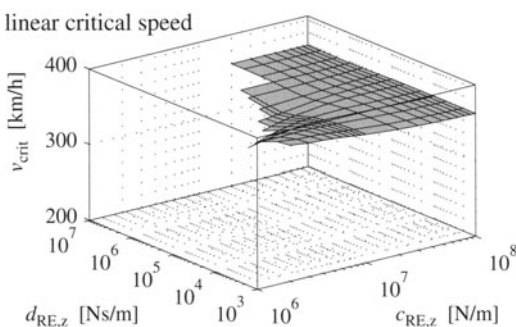
Fig. 16. Variation of vertical elasticity, specific damping

to environmental conditions like the temperature. Nevertheless, low values should be chosen for the radial damping coefficients. The choice of lateral damping coefficients is less critical in case of relatively high lateral stiffness.

6   Summary

A multibody model of a passenger coach with conventional as well as with radial- and lateralelastic wheels is analyzed. Deterministic and stochastic excitations are applied to the wheelsets. The discussion of deterministic excitations turns out that only polygonalizations of order 3$^{\text{rd}}$ and 4$^{\text{th}}$ are able to generate in the droning noise in ICE passenger coaches. The strength of a conventional but polygonalized wheelset is investigated where the wheelset is modeled as flexible body. Vehicle speeds are detected where resonance phenomena occur. The evaluation of dynamic stresses at these speeds turns out that the strength of a conventional wheelset is satisfactory even in case of polygonalization.

The rim and disc of radial- and lateralelastic wheels are modeled as rigid bodies, the elasticity of the so-called third suspension is represented by discrete springs and dampers. General and specific damping of the wheel rim

is investigated and discussed. An optimization of the vertical dynamics leads to reduced droning noise as well as lower force level between wheel and rail. The obtained parameters are applied for investigation of the stability of the wheelset motion. Stability tests, especially with respect to the so-called hunting motion, are performed for various parameter sets. The results turn out that the lateral stiffness must be considerably higher than the radial one to guarantee the stability of the hunting motion. Furthermore, the general damping is a good approximation in case of high lateral stiffness and low radial damping, and may simplify stability analyses in this parameter range. Radial- and lateralelastic wheels with increased bending stiffness and improved parameters are feasible and the recommended parameters may be used for further developments.

References

1. Claus, H.; Schiehlen W. (1998) Modeling and Simulation of Railway Bogie Structural Vibrations. Vehicle System Dynamics Supplement **28**, 538–552
2. Claus, H. (1999) Systemdynamik radialelastischer Räder, ZB-117. Institute B of Mechanics, Stuttgart
3. Claus, H. (1999) Theoretische Lebensdauerabschätzung am Beispiel eines Eisenbahndrehgestells. ZAMM **79**, 493–494
4. Claus, H. (to appear) On Dynamics of Radialelastic Railway Wheelsets. In: Proc. VSDIA'2000 (Budapest, 6–8 November 2000)
5. Claus, H. (to appear) Systemdynamik radialelastischer Räder. In: Proc. GAMM'2001 (Zürich, 12–15 February 2001)
6. Claus, H. (2001) A Deformation Approach to Stress Analysis in Flexible Multibody Systems. Multibody System Dynamics **6**(2), 143–161
7. Claus, H.; Schiehlen, W. (to appear) Stability Analysis of Railways with Radialelastic Wheelsets. In: Proc. IAVSD'2001 (Copenhagen, 20–24 August 2001)
8. Claus, H.; Schiehlen, W. (2002) Symbolic-Numeric Analysis of Flexible Multibody Systems. Mechanics of Structures & Machines **30**(1), 1–30
9. Craig, R. R. (1981) Structural Dynamics, An Introduction to Computer Methods. Wiley, New York
10. Eberhard, P. (1996) Zur Mehrkriterienoptimierung von Mehrkörpersystemen, VDI-Fortschrittbericht **11**(227). VDI Verlag, Düsseldorf
11. Geist at all (1995) Parallel Virtual Machine, A Users' Guide and Tutorial for Networked Parallel Computing. MIT Press, London
12. Guyan, R. J. (1965) Reduction of Stiffness and Mass Matrices. AIAA-J. **2**
13. Hairer, E.; Nørsett, S. P.; Wanner, G. (1987) Solving Ordinary Differential Equations I, Nonstiff Problems. Springer, Berlin
14. Hirsch, T. (2000) Zur Dynamik eines ICE-Mittelwagens, STUD-180 (Claus/Schiehlen). Institute B of Mechanics, Stuttgart
15. Hu, B.; Schiehlen, W. (1997) On the Simulation of Stochastic Processes by Spectral Representation. Probl. Eng. Mech. **12**(2), 105–113

16. Kreuzer, E.; Leister, G. (1991) Programmsystem NEWEUL'90, AN-24. Institute B of Mechanics, Stuttgart
17. Kübler, L. (2000) Parallelisierung einer Evolutionsstrategie für die Eisenbahndynamik, DIPL-84 (Claus/Dignath/Eberhard/Schiehlen). Institute B of Mechanics, Stuttgart
18. Leister, G. (1991) Programmpaket NEWSIM, AN-25. Institut B of Mechanics, Stuttgart
19. Lippold, F. (2000) Dynamische Spannungsanalyse eines ICE-Radsatzes, STUD-192 (Claus/Schiehlen). Institute B of Mechanics, Stuttgart
20. Meinders, T. (1998) Modeling of a Railway Wheelset as a Rotating Elastic Multibody System. Mach. Dyn. Probl. **20**, 209–219
21. Meinders, T.; Meinke, P. (2002) Rotor Dynamics and Irregular Wear of Elastic Wheelsets. Also published in this volume
22. Melzer, F. (1994) Symbolisch-numerische Modellierung elastischer Mehrkörpersysteme mit Anwendung auf rechnerische Lebensdauervorhersagen, VDI-Fortschrittbericht **20**(139). VDI Verlag, Düsseldorf
23. Müller, P. C.; Schiehlen, W. (1976) Lineare Schwingungen. Akademische Verlagsgesellschaft, Wiesbaden
24. Nayfeh, A. H.; Balachandran, B. (1995) Applied Nonlinear Dynamics. Wiley, New York
25. N. N. (1992) Gummigefedertes Schienenrad, Patent DE 33 28 321 C2 (Krupp-Klöckner GmbH). Bundesdruckerei, Berlin
26. N. N. (2000) ANSYS User's Manual. Ansys Inc., Houston, Pennsylvania
27. N. N. (2000) Using Matlab, Version 5.0. The MATH WORKS INC., Natick, Massachussetts
28. Oberle, D. (2001) Zur Laufdynamik von Eisenbahnfahrzeugen mit lateralelastischen Rädern, STUD-198 (Claus/Schiehlen). Institute B of Mechanics, Stuttgart
29. Pallgen, G. (1998) Unrunde Räder an Eisenbahnfahrzeugen. Eisenbahningenieur **49**(1), 56–60
30. Pederzolli, N. (2000) A Multiaxial Fatigue Method and its Application to a Radialelastic Railway Wheel, ZB-119 (Claus/Schiehlen). Institute B of Mechanics, Stuttgart
31. Peić, M. (2000) Transiente Spannungssimulationen mit modalen Spannungsmatrizen, STUD-185 (Claus/Schiehlen). Stuttgart, Institute B of Mechanics
32. Popp, U.; Schiehlen, W. (1993) Fahrzeugdynamik. Teubner, Stuttgart
33. Schiehlen, W. (1997) Multibody System Dynamics: Roots and Perspectives. Multibody System Dynamics **1**(2), 149–188
34. Schiehlen, W. (1999) Elastisches Eisenbahnrad, Patented Design 299 17935.4. Bundesdruckerei, Berlin
35. Schubert, S. (2000) Unrundheiten und Abhilfemaßnahmen bei Eisenbahnrädern, Technischer Bericht. Deutsche Bahn AG, Minden
36. Seibel, C. (2000) Dynamische Spannungsanalyse — Erweiterung der Deformationsmethode, STUD-181 (Claus/Schiehlen). Stuttgart, Inst. B of Mechanics
37. Shabana, A. A. (1997) Flexible Multibody Dynamics: Review of Past and Recent Developments. Multibody System Dynamics **1**(2), 189–222
38. Wallrapp, O.; Eichberger, A. (2000) FEMBS, An Interface between FEM Codes and MBS Codes. DLR, Oberpfaffenhofen
39. Zacher, M. (1990) Unrunde Räder und Oberbausteifigkeit. Eisenbahntechnische Rundschau **45**(10), 605–610

Distributed Numerical Calculations of Wear in the Wheel-Rail Contact

Kurt Frischmuth and Dirk Langemann

Universität Rostock, FB Mathematik,
Universitätsplatz 1, D-18051 Rostock, Germany

Abstract. We study numerical aspects of wear problems in railway mechanics. A short overview of the needed components – a dynamical vehicle/track model, contact geometry and contact mechanics – is given to demonstrate the complexity of the problem. An algorithm for the integration of the domain evolution equation describing wear in the long time scale is presented and discussed. Solutions for simple choices of the model components are studied. Finally, ways to couple the most advanced available models via internet programming are demonstrated on the basis of a vehicle model due to Meinke in a co-simulation with our own geometry program.

1   Introduction

We describe wear in the framework of *domain evolution problems* (cf Sethian, [20], and Fig. 1). Related problems are e.g. curvature driven flows, Stefan
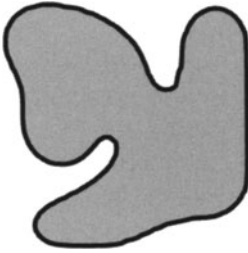
Fig. 1. Example of a domain (used as an extreme test case for assessing integration methods). Under curvature driven motion, $\mathcal{F} = -\kappa$, where κ denotes the curvature of the boundary, the domain should become convex and vanish to a point. Poor numerical methods, however, lead to artificial oscillations of the boundary which in the case of railway mechanics may look like corrugations

problem or the osmotic cell problem. In our case, the speed of the boundary (contact surface) is determined by a *wear law*, the latter using data from a dynamical model as input.

For this paper, the wear process is understood as a pure geometrical change of the body under consideration, i.e. as a shift of the boundary in direction of the inner normal. For now, we do not allow for changes inside the material. Our assumptions lead to the formulation

$$\dot{x}(\tau, t) = \mathcal{F}\big(x(\cdot, t), t\big)\, \boldsymbol{n}\,. \tag{1}$$

Here x is a point on the actual position of the body's surface at time t. The variable τ is used to parameterize the surface.

The speed $\mathcal{F}$ at which the wear surface retreats is determined by a so-called *speed law*. In general, it is not defined by a formula – as it is in the case of curvature driven flows, which we use frequently for testing numerical procedures.

In railway mechanics, to obtain the speed $\mathcal{F}$ at a given position on the actual wear surface, we need to perform a simulation of the motion of the vehicle. During this simulation, given the present surface geometry, the intensity of power dissipation, and hence the wear intensity are determined. Obviously, this is a numerically very expensive procedure, and the result depends on the driving conditions of the simulated vehicle.

The time integration of (1) can be described by the following feedback loop, cf. Fig. 2, in which dynamics, geometry and contact mechanics are coupled together. The numerical solution of a wear problem requires hence two inte-
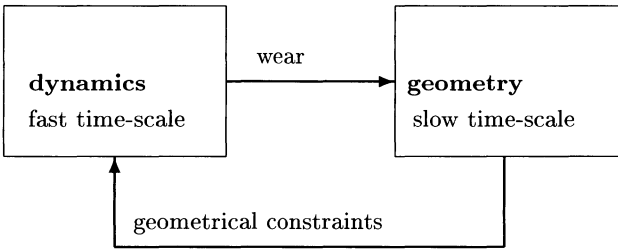
Fig. 2. Knothe's feedbackloop (simplified for autonomous problem)

gration procedures. In an outer loop, solving integration of (1) with respect to the so-called long time scale (or slow time scale), the surface is updated. In an inner loop, simulation with respect to the fast time scale has to be performed. This procedure is – from the point of view of the outer loop – an auxiliary problem, from which we obtain the wear intensity which has the character of a probabilistic density. Here qualitative aspects of the vehicle model, like periodicity, quasi-periodicity or ergodicity, play an important role, [18].

To illustrate the time scales, consider a wheelset rolling at 200 km/h. Under normal conditions, 50 million rotations remove a layer of about 100 μm ... 1 mm. Hence a single-atomic layer of material is abrased in one second, in the same time we observe several recurrences of the lateral motion, 16 revolutions and about 100 oscillations of the normal contact. A reasonable step for the outer time integration would be the order of hours – which would still require about 1000 (quite expensive) steps. For the inner integration, a time step around a millisecond is required to obtain a reasonable accuracy. The ratio is hence of the order of 10^6.

2   Components

In this section we collect some of the ingredients needed to build models capable of describing wear distributed over the circumference of a railway wheel. These components may be combined in different ways to create models for simulations. Our minimal collection may be extended by supplements from other groups.

2.1   Nonlinear Dynamics

We assume that the driving motion is governed by a system of nonlinear equations of motion of the general form

$$m\big(q(t)\big)\ddot{q}(t) = f\big(t, q(t), \dot{q}(t), r(t)\big) . \tag{2}$$

Here we denote by $q \in \mathbb{R}^d$ generalized coordinates, t time, $\dot{}$ time derivative, m the mass matrix and and f generalized forces. External control is expressed
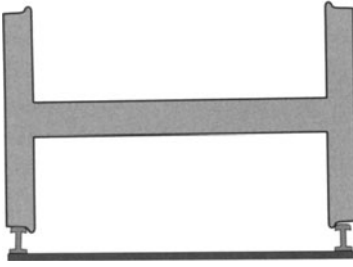
Fig. 3. Wheelset on rails. The evolving surfaces result in complicated relations between lateral shift, elevation and contact point position. For an elastic axle, such relations (dependent on yaw and, in general, also on revolution angle) hold for each of the wheels. For a rigid wheelset, also the rolling angle is determined by a holomorphic constraint

by $r(t)$, which may be a force or a prescribed travelling speed.

The dynamical behavior of a single wheelset, cf. Fig. 3, is already very complex. For low speed, we have a stable trivial solution. Above a critical speed, there is a stable limit cycle. For further increasing speed, there is phase doubling and transition to chaos, cf. for instance [11] and [22].

For the purpose of wear calculation, the dynamical model has to be adapted to the changing geometry in the contact region. This will be discussed in the next two subsections.

2.2   Rolling Contact and Dry Friction

A major component of the dynamical system introduced in the previous subsection is the force vector f on the right-hand side of the equations of motion (2), in particular, the contact forces are difficult to model. For the calculation of friction forces, we have the standard options of the Vermeulen-Johnson law [8], or the Fastsim [3,21] and Contact [9] algorithms. For the long time
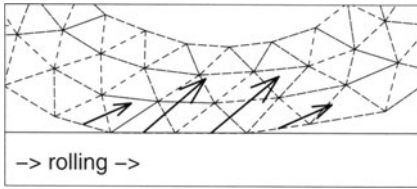
Fig. 4. Disk wheel rolling, accelerating moment and counteracting force F_C. Sparsely discretized 2D-model of a tyre

simulations we need to perform in order to obtain accurate wear speeds, we prefer simpler friction laws in form of algebraic expressions, cf. also [2]. Even a rough approximate solution, cf. Fig. 4, of the underlying Signorini problem, basing on influence functions obtained within the Zastrau/Nackenhorst project, leads to excessive computation times, and is at this moment not possible to carry out for our purposes. In Fig. 5 we present results for the
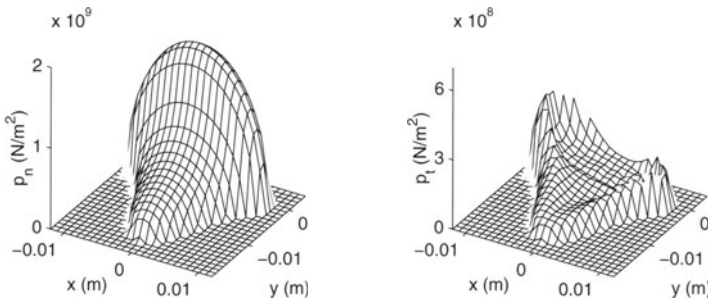
Fig. 5. Contact forces. (a) normal pressure. (b) tangential forces

tractions in the contact zone between wheel and rail for a large normal force. This local distribution of the friction force over the *contact patch* determines how the *dissipated power* has to be mapped to the wear surfaces. In the next subsection, we discuss how to find the location of the *contact point* around which said contact patch is located.

2.3   Geometry

A very sensitive aspect of modeling rolling of realistic wheels on rails is the geometry of the contact, in particular the effective calculation of the geometrical points of contact. Those points are needed for the dynamical model – the calculation of friction forces – and also for the wear model – for assigning the frictional power at a given instant to a place on the contact evolving surface.

For given mbs-coordinates of the bodies, let τ_i^b be the ith surface coordinate of the bth body, $b = 1, 2$, accordingly $n^b = (n_i^b)_{i=1..3}$ the (unique) outer

normal unit vectors to those surfaces x^1, x^2. Then necessary conditions for contact may by formulated as

$$x^1(\tau^1) + \lambda^1 n^1(\tau^1) = x^2(\tau^2)\,, \tag{3}$$

$$x^2(\tau^2) + \lambda^2 n^2(\tau^2) = x^1(\tau^1)\,. \tag{4}$$

This system implies that the normal directions coincide, and that shifting the bodies rigidly in that common normal direction, brings the points $x^1(\tau^1)$ and $x^2(\tau^2)$ to the same spot. This system a is 6×6 nonlinear problem, and for general non-convex bodies hard to solve.

For perfect wheels (axial-symmetric) and rails (prismatic), the system can be simplified considerably, cf. [1]. The solution can be determined from a scalar minimization problem, cf. Fig. 6. Even so, due to the non-convexity, a
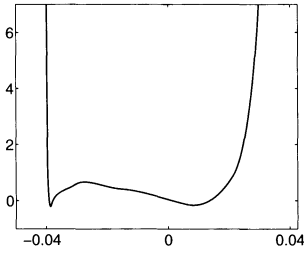
Fig. 6. Cut through the distance function. The minimization of the distance function between wheel and rail surfaces is a non-convex problem, difficult for the existence of more than one local minimum. The distance function depends continuously on the mbs-coordinate q, the minimizer, however, may jump

very accurate starting value for the minimizer has to be found before a fast iterative method (Newton-type) can be successfully applied.

For worn surfaces, reduction to a scalar problem is in general not possible. An implementation of a reasonably quick solver for (3)-(4) is due to Hänler, [9].

For a prismatic rail, however, the simplification still works. Considerable reduction of numerical effort comes from the fact that the first component of the rail's normal vanishes everywhere.

Let us label the wheel by superscript 1, rail by 2, axial coordinate on the wheel by τ^1_1, angular by τ^1_2, and define a function $\tau^1(\tau^1_1) = (\tau^1_1, \tau^1_2(\tau^1_1))$ by the condition

$$\chi(n^1(\tau^1_1, \tau^1_2(\tau^1_1)), q)e_1 = 0\,. \tag{5}$$

Here e_1 points in the direction along the rail, e_3 is vertical, and $\chi(x, q)$ is the actual position of wheel particle x under mbs-coordinates q. Then the parameter of the contact point on the wheel surface is defined via

$$\tau^1_{1\,c}(q) = \operatorname{argmin}\big(\chi(x^1(\tau^1(\tau^1_1)), q) - x^2(x^2_2(\chi(\tau^1(\tau^1_1), q)))\big)e_3\,. \tag{6}$$

The contact point is hence $x^1(\tau^1(\tau^1_{1\,c}))$. Here we assumed that for the rail surface, we have a parametrization of the form $x^2 = x^2(x^2_2)$, i.e. τ^2_1 is along the rail and does not enter the equation, $\tau^2_2 \equiv x^2_2$.

It is obvious that a fast evaluation of the above contact conditions is essential. However, there is no point in tuning procedures for the original profiles, e.g. S1002/UIC 60, since the profiles change essentially during wear simulation. To make this clear, compare the location of the contact point in dependence on lateral shift, once calculated for virgin profiles, once for slightly worn ones, Fig. 7. Note that the contact geometry is the main interface point where the
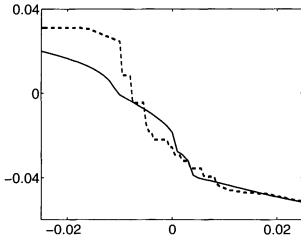
Fig. 7. The axial position of the point of geometrical contact on the wheel surface, for vanishing yaw and roll angles. The continuous line is for the virgin profile pair S1002/UIC60. The dashed line corresponds to slightly worn wheel profile, rail profile still unchanged

feedback loop closes, Fig. 2. Dynamical models which use force models basing on frozen profiles can't work beyond the mere onset of wear.

2.4   Wear Laws

Due to a simple assumption, frequently quoted as Archard's law, cf. [5], for abrasive wear $\mathcal{F}$ depends on the frictional power density dissipated at $x(\tau, t)$. Hence the term $\mathcal{F}$ in (1) is obtained by running the dynamical model (2), and averaging a function of the dissipated power, which is then mapped to the surface of the body in order to obtain a wear distribution.

For example, [19] uses a piecewise constant function for mild and strong abrasion, respectively.

For a comprehensive discussion of factors contributing to wear, we refer to [16].

Observations that reshaped rails wear faster than new ones motivated assumptions about *state variables* assigned to the surface [6].

For this paper, we want to focus on numerical aspects of wear simulation, rather than on sophistication of the wear model. Thus we assume proportionality between worn mass and dissipated energy, which results finally in an equation of the form

$$\mathcal{F} = -\beta s \sigma_t \tag{7}$$

where s is the local creepage, σ_t tangential stress, and β a positive constant. For calculations on the long time scale, on the right-hand side the mean value over a sufficiently long interval of the fast time scale has to be taken.

In terms of Meinder's model, our assumption means that during the exploitation of a given vehicle, we do not change regimes between mild and strong abrasion.

Given this assumption, by matching data on wear of ICE-wheels, the value of the single wear constant β can be found to be of the order 10^{-12} if standard units are used. We prefer, however, to scale the long time scale to 1, and hence we may assume $\beta = 5 \cdot 10^{-6}$.

3   Models

It was Brommundt's original idea, cf. [5] to study evolution of wear in the long time scale on the basis of extremely simple dynamical models in the short time scale. This allowed the application of asymptotic expansions and gave first results very quickly. Despite the fact that we have at our disposal a palette of much more advanced vehicle models, we chose for this paper to follow this approach. It allows us to compare the effects of different analytical and numerical methods on one and the same model. We will see that poor approximations may contribute more to the output than the model itself.

As a **benchmark** model, we propose a rolling disk, driven by a constant moment against a quadratic drag force. The equations of motion for the three degrees of freedom (longitudinal and vertical displacement, angle of rotation) are given by

$$m\ddot{q}_1 = F_R - c|\dot{q}_1|\dot{q}_1 \,, \tag{8}$$

$$m\ddot{q}_2 = F_G + mg + F_N \,, \tag{9}$$

$$J\ddot{q}_3 = M + r(\tau)F_R \tag{10}$$

with

$$\tau = -q_3 \,, \tag{11}$$

$$s = \dot{q}_1 + \dot{q}_3 r(\tau) \,, \tag{12}$$

$$F_N = k\big(r(\tau) - q_2\big)^{3/2} - d\dot{q}_2 \,, \tag{13}$$

$$F_R = -\mu F_N s \,. \tag{14}$$

Equation (11) plays the role of a *contact geometry module*, (12-14) substitute

k $= 1.6742 \cdot 10^9 \,\mathrm{N/m}^{\frac{2}{3}}$
d $= 100 \,\mathrm{Ns/m}$
μ $= 0.2 \,\mathrm{s/m}$
m $= 300 \,\mathrm{kg}$
J $= 37.5 \,\mathrm{kgm}^2$
c $= 0.5 \,\mathrm{Ns}^2/\mathrm{m}^2$
$F_G = -50000 \,\mathrm{N}$
g $= -9.81 \,\mathrm{m/s}^2$
M $= -225 \,\mathrm{Nm}$
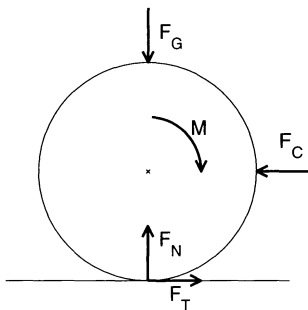
Fig. 8. Disk wheel as a benchmark model

a *contact force model*, the term containing the constant c in (8) simulates the *control motion*. The constants given in Fig. 8 lead to a steady state with longitudinal speed of around 30 m/s, and a frictional power around 19 W. For given radius function $r(\cdot)$, the mean value of frictional power tends to a limit, multiplication with the wear constant β gives finally the evolution speed $\mathcal{F}$:

$$r_t(\tau, t) = -\beta \overline{sF_R} \tag{15}$$
$$\beta = 5 \cdot 10^{-6} \tag{16}$$

where overline stands for taking the temporal mean value, and time is measured in a characteristic time for the life cycle of a wheel. For this case, the approximation of the outer normal by the position vector in (15) is obviously acceptable.

The initial geometry imperfection – a single trough – is defined by B–splines, the data can be downloaded from our website.[1]
In Brommundt's original paper, the dynamical model was completely linear, and the normal force was assumed to be constant. The only nonlinear relation was (5.1). The out-of-roundness is modeled by a sequence of Fourier coefficients. Under these assumptions, every harmonic excites an oscillation of slip and friction force with the same frequency, higher harmonics are caused by the non-linear abrasion hypothesis.
Similar results were obtained by use of approximations to $\mathcal{F}$ in terms of Taylor coefficients of the actual wear surface, [6]. Those lead to evolution equations of the general form

$$r_t = F(u, \nabla r, \nabla^2 r, ...) \tag{17}$$

which for suitable functions F yield speed dependent growth or decay of wavyness of the surface.

4   Numerics

For more realistic models of wear in railway mechanics, we have typically two layers of numerical calculations. In an inner loop, the equations of motion of the dynamical model (vehicle, track) have to be solved. Along with the solution, data essential for wear have to be collected and processed. We call this most essential ingredient of a coupled wear model the *wear collecting algorithm*. From this inner procedure we obtain the speed function $\mathcal{F}$ in dependence on the momentary geometry (together with the control motion). Now, the speed function is passed to the outer time-integration loop in which the wear patterns finally evolve.

---

[1] $r = 0.5$ m everywhere except on an arch of length 0.2 m where it is reduced by 75 µm. Details on `http://alf.math.uni-rostock.de/~kurt/DFG/benchmark/`

A central problem here is accuracy. It turns out, that already for very simple models, it is hard to obtain an accurate speed function. One limiting factor is computation time, because we need trajectories of the dynamical system with rather high precision, but also for a reasonably long time span.

Once the problem of obtaining a good approximation for $\mathcal{F}$, the outer integration loop is straightforward. Since evaluations of $\mathcal{F}$ are extremely costly, we do not recommend too much sophistication. However, as examples show, the common practice of simply applying Euler's explicit method is not the best choice.

Here we need some denotations which we introduce below.

4.1   Discretization

Discretization of the functional variable x and separation leads to an ODE problem of the form

$$\dot{y} = g(t, y, u) \,, \tag{18}$$
$$\dot{u} = h(t, y, u) \,. \tag{19}$$

Here we denote by $y \in \mathbb{R}^{2d}$ the pair of vectors $(q, \dot{q})$ describing the dynamical state of the driving model, while $u \in \mathbb{R}^n$ is the vector of nodal values of a suitable presentation of the change in surface positions on a given grid.

The functions g and h on the right-hand sides are derived from the forces f and the speed function $\mathcal{F}$, cf. (2) and (5.1), respectively.

Now, instead of solving the system (18)–(19) in a straightforward way on a given time interval $[0, T]$ with huge T (not desirable), or solving a modified system with h replaced by ah, a – big, on $[0, T/a]$ instead (common, but dangerous, cf. [17]), we apply the algorithm from [7] for integrating the coupled system.

step 1:   On the time interval $[t_i, t_i + \Delta t_i]$ (which is very long in terms of the y-variables but very short in terms of the u-variables) we integrate the y-equations with frozen u

$$\bar{y}(t) = y(t_i) + \int_{t_i}^{t} g(t, y(t), u(t_i)) \, dt \,. \tag{20}$$

Thus, this low-dimensional part of the system is uncoupled and can be integrated by any suitable method.

step 2:   An approximation for $u(t_{i+1})$, $t_{i+1} = t_i + \Delta t_i$, is calculated by

$$u(t_{i+1}) = u(t_i) + \int_{t_i}^{t_{i+1}} h(t, \bar{y}(t), u(t_i)) \, dt \tag{21}$$

where we use $\bar{y}$ instead of the exact solution.

For the efficiency and accuracy of this algorithm, it is crucial how we approximate the integral on the right-hand side. Using the property that

$$\Delta t_i^{-1} \int\limits_{t_i}^{t_i + \Delta t_i} h(t, \bar{y}(t), u(t_i))\, dt \tag{22}$$

approaches a limit very quickly, we may substitute the integral over the interval $[t_i, t_i + \Delta t_i]$ by

$$\frac{\Delta t_i}{\Delta t_{sim}} \int\limits_{t_i}^{t_i + \Delta t_{sim}} h(t, \bar{y}(t), u(t_i))\, dt\,, \tag{23}$$

where Δt_{sim} is a reasonable time span for simulation of the y-system. The latter is chosen by monitoring the limit of the integral mean value, terminating when changes become negligible.

In the sequel, given the interpretation of (18) and (19), we will refer to step 1 of the algorithm as *wear collection*, and to step 2 as *geometry updating*.

4.2   Distributed Calculations

The algorithm described in the previous subsection is not well suited for *parallel computations*, it is essentially *sequential*. Nonetheless, for several practical reasons, it is sensitive to perform some routines on different computers. In fact, often more advanced models are not easy to port from their original environment, hence it is far easier to couple them while leaving them where they were developed. (Usually, they are still being developed.)

Thus we agreed with other groups of the DFG Priority Programme 1015 (Meinke/Meinders, Stuttgart, [19] and Popp/Kaiser, Hannover, [12]) on coupling our models by the most simple possible interface – which is on the other hand the safest to work with. Anticipating that all partial models have a considerable cost in terms of CPU-time, it is clear that time for the exchange of data is not the limiting factor. All results from one model are written on ASCII text files and then transported by FTP to the remote machine. Timing is done by waiting for needed input files. For all tests, losses due to communication time were minimal.

We implemented distributed calculations with Meinke/Meinders the following way. A (simple) vehicle model used as driving component is integrated on one CPU (in Stuttgart). Along the trajectories dissipated power is calculated and projected onto the surface grid by a method using geometrical data on the contact geometry. The resulting wear intensity is sent to our server, where the surface is updated, new contact geometry data are calculated and sent to (the meantime idle) machine in Stuttgart.

To make this loop really closed, the dynamical model has to depend on the

contact geometry. In our case, this is the case due to dependencies of forces and moments on the position of the contact point and the curvatures of the profiles there. Further, the wear collection is very sensitive to the contact point position, cf. Fig. 7.

It has to be mentioned that for the simulation of realistic wear processes the number of outer loop time steps is not excessive, it rarely exceeds 100.

Total time needed for the file transfer is measured in seconds, while the overall computation time is – even for simple components – the order of hours. Consequently, we have practically no loss of efficiency if compared with running the simulation on one computer. What we gain is that we save the effort to get always the actual version of the software developed in the other group (or groups), together with all the formal problems that arise from that (e.g. obtaining licences).

4.3 Testing the Speed Function

Before approaching the very time consuming integration of the shape evolution equation (outer loop), it is sensitive to evaluate and discuss the speed function $\mathcal{F}$, i.e. to test step 1 of the algorithm, wear collection, without geometry update. We perform all tests on the initial radius described in Sec. 3. There are several sources of errors (from our own experience), which we want to point out. All have catastrophal impact on the results.

The typical errors come from (i) too small accuracy of the time integration of (18), (ii) bad discretization of the manifold x and (iii) too small Δt_{sim} in the algorithm.

For (i) we remark that matlab's built-in solvers, with standard accuracy options, perform poorly on the benchmark problem. In order to get a feeling for the model, it is helpful to monitor the vertical motion, presented by Fig. 9. The trajectory converges to the presented cycle, the 'shortcut' caused
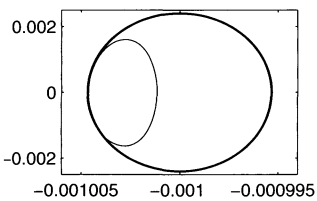
Fig. 9. Integrating the benchmark model. Vertical phase as signature of the dynamical model

by the initial imperfection is taken once in eight rounds. As illustration for (ii), we compare the correct result, confirmed by several methods, with erroneous wear speeds, Fig. 10. The source of the error here is a typical effect for trigonometric polynomials – on localized defects, like a trough, they tend to oscillations and miss the extrema. From (c) we infer that even the best approximation by a trigonometric polynomial is not suited for reliable wear
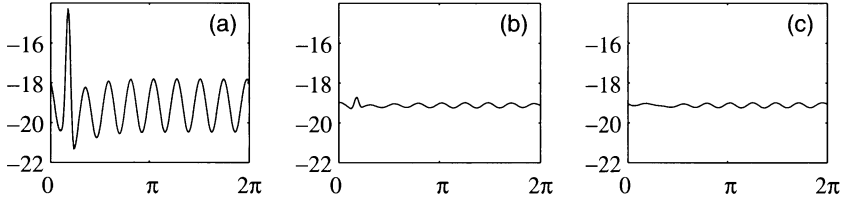
Fig. 10. (a) wear intensity for benchmark model, (b) same with Fourier approximated radius, (c) Fourier approximation of (b)

simulations. A comparison with models without dynamical response in ver-
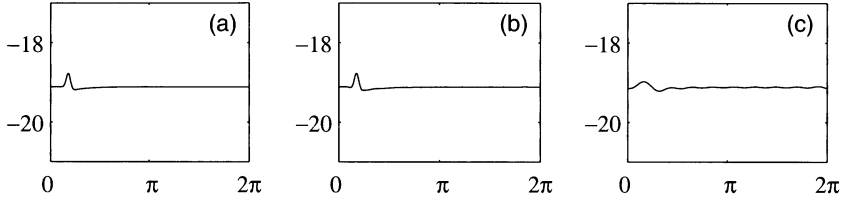
Fig. 11. (a) wear intensity in model with frozen normal force, (b) same with Fourier approximated radius, (c) Fourier approximation of (b)

tical direction – i.e. with normal frozen force, equal to gravity plus external load, shows large differences. In Figs 10 and 11, we used the same order of trigonometric polynomials $N = 10$ as in [5], i.e. 21 coefficients. Our results show the importance of the vertical components of the dynamical model, and discourage the use of trigonometric polynomials.

Let us visualize the third source of errors in a modified version of the benchmark model – we superimpose a lateral motion according to a linear model

$$m\ddot{q}_4 = -k_4 q_4 \qquad \text{with } k_4 = 4147\,\text{N/m} \tag{24}$$

A typical run of the algorithm collecting data on the wear distribution yields the following sequence of figures, Fig. 12 It is obvious from Fig. 12 that too quick termination of the wear collection may lead to possibly most spectacular, but very wrong results. Note that for Fig. 12 (a) a different scale had to be used.

5   Results

In this section we present several results obtained for the integration of the surface evolution equation (1) with the speed function $\mathcal{F}$ calculated by the algorithm from Sec. 4. As a rule, dynamical simulations (inner loop) were
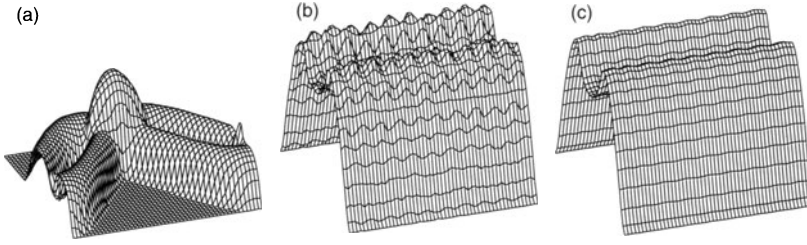
Fig. 12. Stages of the wear collection. **(a)** early. **(b)** medium. **(c)** final

performed on a SUN workstation with 1.2 GHz processor, geometry and graphics on a PC with and AMD Athlon processor at 1.6 GHz, data files were exchanged by was done by FTP.

The examples relate to the questions in how far wear patterns are determined by initial imperfections, and whether they stay in place or travel.

For the benchmark model from Sec. 3, a time interval of length 4.0 (corresponds to ca. 1000 h) gives reasonable wear patterns. The results are pretty robust with respect to parameters of the numerical method, e.g. the time step of the outer loop or the use of predictor/corrector steps.

We obtain always a pattern with 8 maxima and minima, as indicated by Fig. 10 (a). With the evolution, the wear speed tends almost to a sinusoidal distribution.

For higher speed, the number of minima decreases. However, the variation of the radius may grow considerably faster, so that we get lift off very quickly, [3]. This violates the model assumptions, hence we stop simulations at the first occurrence of zero normal force. For a three times higher speed, we obtain 3 humps, but life time (time to first bumping of the wheel) decreases 8 times, cf. Fig. 13. We observe further that the frictional power is phase-
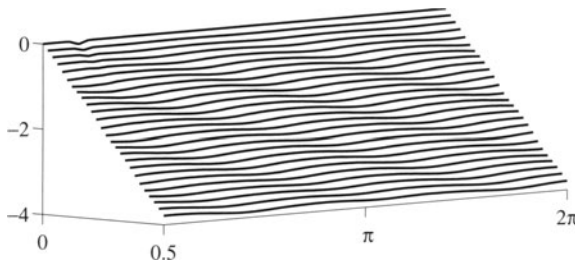
Fig. 13. Change of radius (in mm) vs angle and time (in 10^6 s), for a moment $M = -2000$ Nm. For the original benchmark, we observe a pattern of order 8. The higher moment results in a speed of ca. 320 km/h and only 3 humps

shifted with respect to the radius, hence the wear pattern is not fixed but moves around against rolling direction.

This effect may not be observed for cases where the variation growth too fast.

Than lift off takes place before the shift of the pattern becomes significant. Further, a preexisting imperfection, like a flat spot, 'pins' the pattern to the surface until it is levelled out.

We have to point out that for this presentation, we chose a driving wheel. The mentioned effects occur likewise for a wheel coupled elastically to a perfect partner wheel on a wheelset, cf. [5].

In conclude this section by some remarks on randomizing some of the model parameters.

5.1   Randomized Control

Following a hint from a referees, we have performed calculations with randomized data. Among others, we have run a vehicle over randomized ground.

For a random geometric error imposed on the vertical position of the rail head, we obtain – in accordance with a lemma from [17] – that the results coincide with those for a slightly changed wear law. For a given out-of-the-
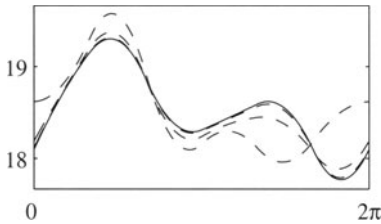
Fig. 14. 1D wear simulation, disk wheel model on wavy ground. The *solid line* shows the power dissipation (W) for an ideal track, the *dotted lines* indicate mean values over increasing rolling times (1,10,100 (s)) on randomized track

round wheel, on perfect ground, we would obtain the creepage distribution indicated in Fig. 14, solid line.

On wavy ground, we obtain a position (= time) dependent creepage. However, the mean value curves converge (with running time going to infinity) to the output of the undisturbed ground case (dotted curves). In this case, we have an *averaging effect* as should be expected on the basis of [17].

Analogously, we introduced a random lateral force f_4 acting on a damped variant of the lateral dynamic (24)

$$m\ddot{q}_4 = -k_4 q_4 - d_4 \dot{q}_4 + f_4(t) \qquad \text{with } d_4 = 514\,\text{Ns/m}. \tag{25}$$

The parameters of the the random force distribution f_4 determine the shape of the final geometry, Fig. 15. Note that the damping has considerable influence on the distribution along the wheel axis.

6   Conclusions

It is as desirable as unrealistic to have the most advanced models of all sub-systems involved in the feedback-loop of system dynamics and long-term
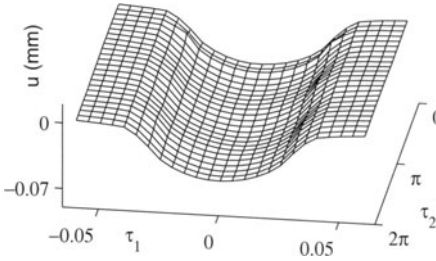
Fig. 15. 2D extension of benchmark problem. The force $f_4(t)$ is a time series of random impulses with zero mean and a deviation of 3000 N

behavior, and run them in one single program or package. Instead, we use simplifications where possible, and couple components of other groups via internet programming. It was shown in Sec. 3 that for studying polygonalisation, it is essential to include vertical dynamics in a sensible way. That way results depend reasonably on data like initial surface deviations and speed of control motion, and allow a consistent interpretation.

From the point of view of long-term changes, models of vehicles, track and subgrade are components that determine the right-hand side of an evolution problem for the domain occupied by the wearing parts. Proper methods of integration for domain evolution problems are as important as fast methods for solving the equations of motion of the dynamical systems. Benchmark problems designed for testing the long-term integration have been formulated. For closing the feedback loop, the dynamical system has to respond to changes of geometry.

In this project, we have mainly studied wear patterns on wheels; patterns on rails have been omitted to keep this paper within the given limits. For the same reason, temperature effects on friction and wear laws have not been discussed, [15]. In the future, roughness, the real contact areas and local forces, and in turn the changes of surface texture, should be modeled and included into the feedback loop.

7   Acknowledgements

References

1. Arnold, M., Netter, H. (1998) Approximation of contact geometry in the dynamical simulation of wheel-rail systems. Mathematical and Computer Modelling of Dynamical Systems **4**, 162–184

2. Bogacz, R., Frischmuth, K. (2001) Slip waves on curved track. In: Proceedings of the 7th German-Polish Workshop on Dynamical Problems in Mechanical Systems, Osieki

3. Bogacz, R., Kowalska, Z. (2001) Computer simulation of the interaction between a wheel and a corrugated rail. Eur. J. Mech. A/Solids **20**, 673–684

4. Brommundt, E. (1996) A simple mechanism for the polygonalization of railway wheels by wear. Mechanics Research Communications Basic and Applied

5. Chudzikiewicz, A. (2000) Zużycie kół kolejowych. In: Proceedings of the 7th Conference on Computer Simulation in Research and Development, Kościelisko, September 14–26

6. Frischmuth, K. (2000) Wear models with internal state parameters. Machine Dynamics Problems, **24**, No. 1, 79–86

7. Frischmuth, K. (2000) Temporal evolution of wear profiles. TransComp, Zakopane

8. Frischmuth, K. (2001) Contact, motion and wear in railway mechanics. Journal of Theoretical and Applied Mechanics, 3, **39**

9. Hänler, M. (1998) Private communication.

10. Johnson, K. L. (1985) Contact mechanics. University Press, Cambridge

11. Kaas-Petersen, Ch. (1986) Chaos in a Railway Bogie. Acta Mechanica **61**, 89–107

12. Kaiser, I., Popp, K. (2000) The behaviour of a railway bogie in the mid-frequency range. Euromech Colloquium 409, Hannover, March 6-9

13. Kalker, J. J. (1990) Three-Dimensional Elastic Bodies in Rolling Contact. Vol. **2** of Solid Mechanics and its Applications, Kluwer Academic Publisher, Dordrecht

14. Kik, W., Piotrowski, J. (1996) A fast, approximate method to calculate normal load at contact between wheel and rail and creep forces during rolling. In: Zobory, I. (Editor): Proc. of the 2nd MiniConference on Contact Mechanics and Wear of Rail/Wheel Systems, TU Budapest, 52–61

15. Knothe, K. (2001) Probleme der Berechnung von Rollreibungskräften. Vortrag im FK Numerik, Universität Rostock, 7. November

16. Kragelski, I. V. (1965) Friction and Wear. Butterworth Washington DC

17. Langemann, D. (1999) Numerische Analyse abrasiv verschleißender mechanischer Systeme. Fortschritt-Berichte 12.**392**, VDI-Verlag Düsseldorf

18. Langemann, D. (2000) Numerical analysis of wear processes. submitted to SIAM Journal of Applied Mathematics

19. Meinders, T., Meinke, P. (2000) Development of polygonalized wheels and their effects on the dynamical system. Euromech Colloquium 409, Hannover, March 6-9

20. Sethian, J. A. (1996) Level set methods. Cambridge University Press, Cambridge

21. Shen, Z. Y., Hedrick, J. K., Elkins, J. A. (1983) A Comparison of Alternative Creep Force Models for Rail Vehicle Dynamic Analysis. In: The Dynamics of Vehicle on Road and on Track, Proc. of 8th IASVD Symposium held at Massachusets Institute of Technology, Cambridge, MA, August 15–19

22. True, H. (1993) Dynamics of a rolling wheelset. Applied Mechanics Reviews, **46**, No. 7

Modeling and Simulation of the Mid-Frequency Behaviour of an Elastic Bogie

Ingo Kaiser and Karl Popp

Institut für Mechanik, Universität Hannover, Appelstraße 11, D-30167, Germany

Abstract. Some effects occurring at modern high speed railway vehicles like grumbling noise and polygonalization of the wheels are located in the mid-frequency range between 50 Hz and 500 Hz. Since multi-body systems consisting of rigid bodies are not sufficient for this frequency range, a description of the rotating elastic wheelsets is developed in this paper. This description results in linear differential equations with constant coefficients and provides a very simple investigation of the vehicle and an easy integration into a complete vehicle-track model.
Furthermore, the behaviour of the complete system consisting of the vehicle with elastic wheelsets, the wheel-rail contact, and an elastic track, is simulated. The influence of the wheelset elasticity is investigated for the cases of transient motion and limit cycle behaviour. Moreover, the influence of the car body is investigated. As the polygonalization is an important phenomenon, the vibrations excited by an unround wheel are calculated.

1   Introduction

Modern high-speed railway vehicles show some phenomena, which are not fully understood. The probably best known phenomenon is the grumbling noise, located in the range around 100 Hz. Furthermore, at vehicles, which show this grumbling noise, the effect of wheel polygonalization occurs, i.e. non-uniform wear over the circumference of the wheels, resulting in unround wheels. Not only the vehicles, but also the track suffers from damage, for example pulverization of the ballast. More informations concerning these phenomena can be obtained from [1] and [2].
These effects are located in the mid-frequency range, i.e. between 50 Hz and 500 Hz. To describe these phenomena, a "conventional" modeling of the vehicle as a multi-body system (MBS) consisting of rigid bodies is not sufficient. Because the mentioned effects, especially the polygonalization, are mainly caused by the wheel-rail contact, the elasticity of the wheelsets shall be taken into account, because they are the nearest components to the contact. Concerning the wheelset elasticity, the following questions are tried to be answered:

1. Which description provides an efficient modeling of the elastic wheelsets?
2. Which influence does the wheelset elasticity have?
3. What is the influence of the components next to the wheelsets on the dynamic behaviour?

2   Modeling

The complete vehicle-track system can be split up into three subsystems: the vehicle, the wheel-rail contact, and the track. This is shown in Fig.1.
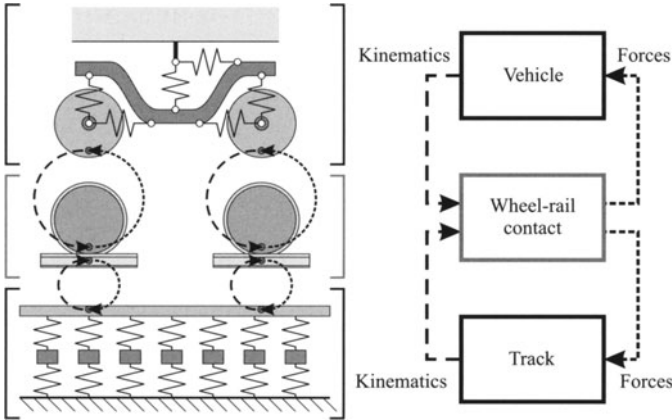
Fig. 1. Subsystems of the vehicle-track system

In view of the vehicle and the track, the forces generated by the wheel-rail contact are regarded as external forces. The excitation of the vehicle and the track leads to motions. The corresponding kinematic quantities are fed back to the wheel-rail contact. The interchange of the kinematics and the forces between the wheel-rail contact and the vehicle respectively the track is performed at nodes, where the subsystems are coupled.

2.1   Modeling of the Vehicle

The vehicle consists of one car body and two bogies, cp. Fig.2. Each bogie consists of the bogie frame, the bolster, and two wheelsets. The bolsters are connected to the car body by a pivot, their only degree of freedom is the yaw motion ψ. For all other bodies, all six degrees of freedom are permissible. All bodies are connected by linear springs and dampers, only the yaw damping between the bolsters and the car body uses dry friction.

As said before, the elasticity of the wheelsets is to be taken into account, so the vehicle is modeled as an Elastic Multi-Body System (EMBS). The motions resulting from the deformation are superposed to the motions of the undeformed body, which are called "rigid body motions" in the following. The consideration of the deformation is performed by modal synthesis. The modes used for the modal synthesis are calculated using a model for the structural dynamics like a finite element model. Usually in modal synthesis
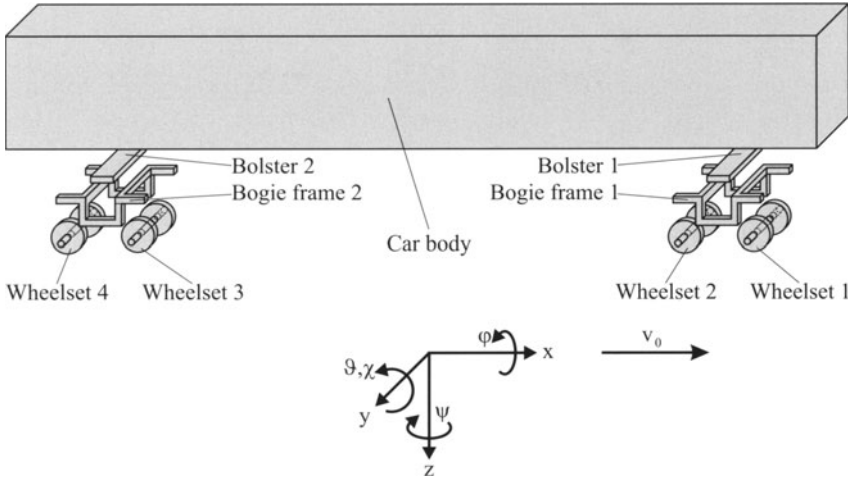
Fig. 2. Components of the vehicle

the question arises, how many modes are necessary to perform a sufficient description of the behaviour. This will be discussed later.

The model for the vehicle should fulfill the following requirements:

1. The model should have a simple mathematical structure.
2. The large rotation of the wheelset should be taken into account.
3. The formulation of the EMBS should be suitable for arbitrary modeling of the structural dynamics of the wheelset.

Usually, the modes for the description of the deformations are given in a body-fixed coordinate system. However, the problem is, that the wheelsets perform large rotations related to the inertial system, which cannot be linearized. This large rotations lead to nonlinear equations or linear equations with periodic coefficients, if the modal synthesis is performed in the coordinate system fixed to the wheelsets.

This problem can be solved by using a special sequence of the rotations, which transform the inertial system $\mathcal{I}$ into the body-fixed coordinate system $\mathcal{B}$ via an intermediate system $\mathcal{A}$. The rotation around the large angle has to be the last in the sequence, which yields the following rotational matrices

$$\mathbf{S}^{\mathcal{IB}} = \mathbf{S}^{\mathcal{IA}}\mathbf{S}^{\mathcal{AB}} \,,$$

$$\mathbf{S}^{\mathcal{IA}} = \begin{bmatrix} \cos\psi_{WS} & -\cos\varphi_{WS}\sin\psi_{WS} & \sin\varphi_{WS}\sin\psi_{WS} \\ \sin\psi_{WS} & \cos\varphi_{WS}\cos\psi_{WS} & -\cos\varphi_{WS}\sin\psi_{WS} \\ 0 & \sin\varphi_{WS} & \cos\varphi_{WS} \end{bmatrix} \,,$$

$$\mathbf{S}^{\mathcal{AB}} = \begin{bmatrix} \cos\chi_{WS} & 0 & \sin\chi_{WS} \\ 0 & 1 & 0 \\ -\sin\chi_{WS} & 0 & \cos\chi_{WS} \end{bmatrix} \,. \tag{1}$$

Here, the angle of the rolling motion χ_{WS} is large, while the angles ψ_{WS} of the yaw motion and φ_{WS} of the roll motion are small. This sequence provides, that the large rotation is performed around the axis of rotational symmetry of the wheelset. A new coordinate system $\mathcal{A}$ is defined, which performs all motions of the undeformed wheelset except the large rotation.

Because of the rotational symmetry of the wheelset, the modal synthesis is performed in the system $\mathcal{A}$. Due to the rotational symmetry, cylindrical coordinates r, θ and y are used. This yields the following vector describing the position of one point of the wheelset:

$$
\mathbf{r}_{CP}^{\mathcal{B}} = \begin{bmatrix} r\sin\theta \\ y \\ r\cos\theta \end{bmatrix} + \sum_i \mathbf{\Phi}_i(r,\theta,y)q_i(t) \ , \ \ \mathbf{\Phi}_i(r,\theta,y) = \begin{bmatrix} U_i(r,\theta,y) \\ V_i(r,\theta,y) \\ W_i(r,\theta,y) \end{bmatrix} . \quad (2)
$$

Here, $\mathbf{\Phi}_i(r,\theta,y)$ and $q_i(t)$ are the modal functions and the modal coordinates, which describe the deformations.

The coordinate θ is not a material coordinate, but a local one. The relation between the local coordinate θ and the material coordinate ϕ is given by:

$$
\theta = \phi - \Omega_0 t + \sum_k \Theta_k(r,y)s_k(t) . \quad (3)
$$

Here, Ω_0 denotes the reference angular velocity of the wheelset. Deviations from the reference motions are superposed. $\Theta_k(r,y)$ are the modal functions for the torsional motions. This special consideration of the torsion is necessary, because torsional deformations do not change the shape of the wheelset. This yields the following expression for the velocity of the considered point:

$$
\mathbf{v}_{CP}^{\mathcal{B}} = \frac{d\mathbf{r}_{CP}^{\mathcal{B}}}{dt} = \frac{\partial \mathbf{r}_{CP}^{\mathcal{B}}}{\partial t} + \frac{\partial \mathbf{r}_{CP}^{\mathcal{B}}}{\partial \theta}\frac{d\theta}{dt}
$$

$$
\frac{\partial \mathbf{r}_{CP}^{\mathcal{B}}}{\partial t} = \sum_i \mathbf{\Phi}_i(r,\theta,y)\dot{q}_i(t) \ , \ \ \frac{\partial \mathbf{r}_{CP}^{\mathcal{B}}}{\partial \theta} = \begin{bmatrix} r\cos\theta \\ 0 \\ -r\sin\theta \end{bmatrix} + \sum_i \frac{\partial \mathbf{\Phi}_i(r,\theta,y)}{\partial \theta}q_i(t)
$$

$$
\frac{d\theta}{dt} = -\Omega_0 + \sum_k \Theta_k(r,y)\dot{s}_k(t) \ , \ \ |\Omega_0| \gg \left| \sum_k \Theta_k(r,y)\dot{s}_k(t) \right| . \quad (4)
$$

It can clearly be seen, that a large angular speed Ω_0 appears, but no large angle. The rotational displacements between the coordinate system $\mathcal{A}$, where the modal synthesis is performed, and the inertial system $\mathcal{I}$ and the system $\mathcal{F}$ fixed to the bogie frame are small, so that a linearization is possible. The advantages of this description are:

1. The coupling between the elastic wheelsets and the bogie frame is described by linear equations with constant coefficients.
2. The wheel-rail contacts do not move around the wheels.

The equations of motion for the vehicle have the following structure:

$$\mathbf{M}\,\ddot{\mathbf{y}}(t) + (\mathbf{D} + \varOmega_0\mathbf{G})\,\dot{\mathbf{y}}(t)\,\left(\mathbf{K} + \varOmega_0^2\mathbf{Z}\right)\,\mathbf{y}(t) = \mathbf{h}(t)\,. \tag{5}$$

The vector $\mathbf{y}(t)$ contains the generalized coordinates. The matrix $\mathbf{M}$ is the mass matrix, the matrix $\mathbf{D}$ is the damping matrix resulting from the viscous coupling elements between the wheelset, the bogie frame, and the car body, and the matrix $\mathbf{K}$ is the stiffness matrix resulting from the elastic coupling elements and the modal stiffnesses of the wheelsets. The matrices $\varOmega_0\mathbf{G}$ and $\varOmega_0^2\mathbf{Z}$ describe gyroscopic and centrifugal forces, respectively. The vector $\mathbf{h}(t)$ represents generalized external forces, resulting from the wheel-rail contact, from the gravitation, and from the nonlinear yaw damping. Because of the symmetric structure of the vehicle, the equations of motion can be split up into two separate systems for symmetric and antimetric motions.

The very simple mathematical structure of the equations of motion for the vehicle provides a very easy integration into complete vehicle-track models, not only in the time domain, but also in the frequency domain: This vehicle model was integrated into the program system TTI by Kruse, which works in the frequency domain, see [3]. Furthermore, a fast analytical solution of the equations for the vehicle can be performed. This can be used for a convergence check or for parameter studies, which will be shown later.

2.2   Finite Element Model of the Wheelset

As said before, modal functions are needed for the modal synthesis. Here, the eigenfunctions of the free wheelset are used as modal functions. They are calculated using a finite element model of the wheelset. Because of the fact, that no point of the wheelset is fixed at all times, there are no geometrical boundary conditions for the wheelset, so the eigenfunctions of the free wheelset have to be taken. The finite element model used is shown in Fig.3. Here, one can take advantage from the symmetry properties of the wheelset. Because of the symmetry to the middle plane, the modeling of only one half of the wheelset is sufficient. The brake discs and the wheelset bearings are considered as rigid bodies. The axle of the wheelset is considered as a one-dimensional continuum, having the properties of a bar, a torsional rod, and a Rayleigh beam. The logitudinal and torsional deformations are described by linear shape functions; for the bending, cubic shape functions are used. The wheel is considered as a two-dimensional continuum, having the properties of a disc and a Kirchhoff plate. In circumferential direction, trigonometric functions are used, taking advantage from the rotational symmetry of the wheelset. In radial direction, linear and cubic shape functions are used for the disc and the plate, respectively.

The calculated eigenfrequencies of the wheelset are listed in Table 1.
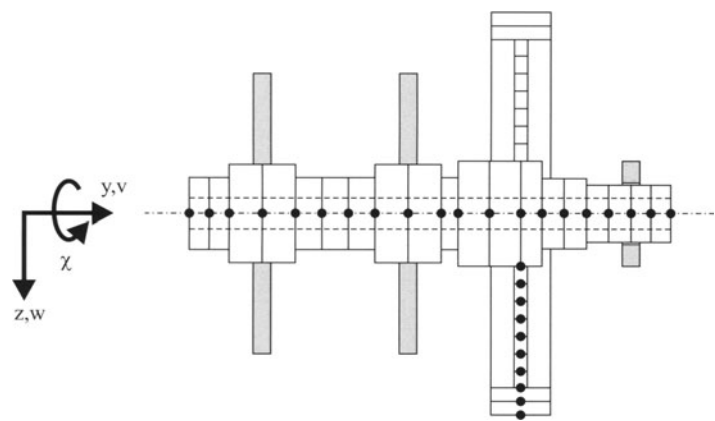
Fig. 3. Finite element model of the wheelset

Table 1. Eigenfrequencies of the wheelset

	symmetric	antimetric
Torsional modes	267 Hz, 597 Hz, 3158 Hz,	82 Hz, 459 Hz, 630 Hz,
	4338 Hz, 6439 Hz, 6860 Hz,	3158 Hz, 4338 Hz, 6439 Hz
	6965 Hz, 7633 Hz, 10038 Hz	6945 Hz, 7633 Hz, 10038 Hz
Umbrella modes	195 Hz, 808 Hz, 1686 Hz	247 Hz, 1523 Hz, 1746 Hz
	2168 Hz, 2978 Hz, 3887 Hz	2539 Hz, 3877 Hz, 4969 Hz
	4976 Hz, 7790 Hz, 10612 Hz	7789 Hz, 10022 Hz
Bending modes	81 Hz, 185 Hz, 500 Hz	131 Hz, 334 Hz, 613 Hz,
	780 Hz, 1262 Hz, 1671 Hz,	1235 Hz, 1669 Hz, 1946 Hz,
	1946 Hz, 2197 Hz, 3024 Hz,	2109 Hz, 2223 Hz, 3024 Hz,
	4370 Hz, 4762 Hz, 7195 Hz,	4370 Hz, 4762 Hz, 7195 Hz,
	7390 Hz, 8998 Hz, 9812 Hz,	7390 Hz, 8998 Hz, 9926 Hz,
	10032 Hz	10106 Hz

2.3   Modeling of the Contact

The input of the wheel-rail module are the kinematical quantities, i.e. the translational and the rotational displacements and velocities of the nodes, where the wheel-rail contact is connected to the wheel and the rail, respectively. The output of the module are the moments and torques acting in the nodes.

The problem is, that the position of the contact between the wheel and the rail varies, but for the EMBS-formulation it would be more advantageous,

if the nodes are fixed to the bodies. To solve this problem, in the contact module, the wheel and the rail are considered as rigid bodies, except the deformations in the contact zone. This assumption seems to be applicable, because the wheel rim and the rail head are comparatively massive parts; they can move due to the deformations of the wheel and the rail, but hardly change their shape. The assumption of rigid bodies provides a very simple calculation of the kinematics for the actual contact position and of the forces and torques acting at the nodes. The coupling of the wheel with the contact module is shown in Fig. 4, the coupling of the contact and the rail is performed in an analogous way.
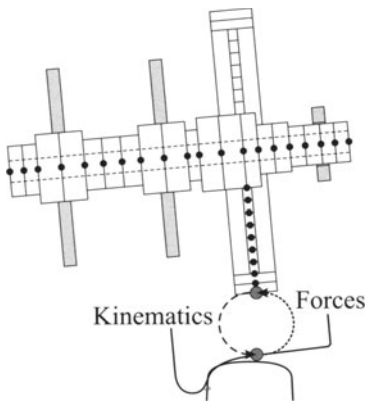
Fig. 4. Coupling of wheel and rail by the contact module

First, the relative displacement of wheel and rail is calculated. For the contact condition in normal direction, a contact spring is assumed, so there is no constraint. A penetration of the undeformed surfaces of the bodies is permissible. So, the contact points are found by the condition, that the penetration vector connecting the contact points has to be orthogonal to the surfaces of the undeformed bodies. The corresponding geometry is illustrated in Fig.5.
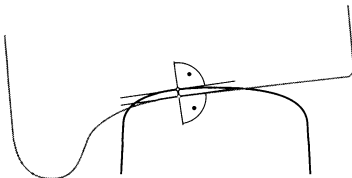
Fig. 5. Contact geometry

The wheel has the profile S 1002, the rail has a slightly modified profile UIC 60; the modification is the replacement of the middle circle of the rail by an

ellipse, according to [8]. This provides a single point contact between the wheel and the rail. After the contact points are known, the penetration of the surfaces and the velocities at the contact point are calculated.

In the second step, the forces and torques are calculated. A Hertzian normal contact is assumed, the necessary curvatures can be calculated from the profiles. The tangential forces are calculated by the algorithm of Polach, given in [4], which bases on the theory of Kalker. This algorithm provides a very fast calculation of the nonlinear contact forces.

2.4   Modeling of the Track

The track model used was developed by Fingberg in [5] and [6]. The model bases on the calculations of Ripke [7], who developed a detailed model of the track. The frequency response functions of the track calculated by Ripke are approximated by linear decoupled mass-spring-damper oscillators.

For the lateral motions of the rails, the equations of motions are given by:

$$m_{y,i}\ddot{q}_{y,i}(t) + b_{y,i}\dot{q}_{y,i}(t) + c_{y,i}q_{y,i}(t) = F_y(t) \ , \ i = 1 \ldots 5,$$

$$v_{Track}(t) = \sum_{i=1}^{5} q_{y,i}(t) \, . \tag{6}$$

Here, $m_{y,i}$, $b_{y,i}$, and $c_{y,i}$ are the parameters for the equivalent mass, damping, and stiffness, respectively. The lateral force and the lateral displacement of the rail are denoted by $F_y(t)$ and $v_{Track}(t)$, respectively.

For the vertical motions of the rails, the influence of the discrete support by the sleepers is taken into account. The equations of motion are given by:

$$m_{z,i}(t)\ddot{q}_{z,i}(t) + b_{z,i}(t)\dot{q}_{z,i}(t) + c_{z,i}(t)q_{z,i}(t) = F_z(t) \ , \ i = 1 \ldots 3,$$

$$w_{Track}(t) = \sum_{i=1}^{3} q_{z,i}(t) \, . \tag{7}$$

The vertical force and the vertical displacement of the rail are denoted with $F_z(t)$ and $w_{Track}(t)$, respectively. The parameters $m_{z,i}(t)$, $b_{z,i}(t)$, and $c_{z,i}(t)$ of the mass, the damping, and the stiffness are time-dependent. As an example, the time dependency of the mass parameter m_i is given by the following equation, the equations for the damping and the stiffness have an analogous structure:

$$m_i(t) = \frac{m_{si} + m_{mi}}{2}$$
$$+ \frac{m_{si} - m_{mi}}{2} \left[\cos\left(\frac{2\pi x_{WS,n}}{l}\right) + \frac{1}{4}\left(1 - \cos\left(\frac{4\pi x_{WS,n}}{l}\right)\right) \right],$$
$$x_{WS,n} = v_0 t + \Delta x_{WS,n} \, . \tag{8}$$

Here, l and $x_{WS,n}$ denote the sleeper spacing and the current position of the wheelset n.

These equations are used for the description of the track elements at each wheel. However, the elements are independent from each other, i.e. there is no direct coupling between two track elements.

3   Simulation Results

To investigate the influence of the wheelset elasticity, simulations are performed using the models described above. The notation of the components corresponds to Fig.2.

If the nonlinear yaw damping of the vehicle is replaced by a linear damping, the frequency response function can be calculated in a very fast way. This is performed to check the convergence for the modal synthesis and to investigate the influence of the gyroscopic effects caused by the wheelset rotation.

Since the complete vehicle-track model contains nonlinear components, mainly the wheel-rail contact, the equations can only be solved by using numerical integration schemes. Here, a Runge-Kutta method of the fourth order has been applied. For the complete vehicle-track system, the following questions are investigated:

1. What is the influence of the discrete rail support?
2. What is the influence of an unround wheel?
3. Which influence does the wheelset elasticity have on the limit cycle behaviour and the transient response?
4. What causes the differences between rigid and elastic wheelsets, if there are any?

3.1   Frequency Response of the Vehicle

In case of an excitation by harmonic forces, the differential equations (5) for the vehicle can be reduced analytically to algebraic equations, so that the frequency response function for the vehicle can be calculated. To check the convergence of the modal synthesis, the frequency response function of the vehicle is calculated with varied numbers of modes. The deviations between the calculated curves give answer to the question, how many modal functions are necessary. As an example, the results of the investigation for symmetric vertical harmonic excitation are shown in Fig.6. The frequency limits for the bending modes correspond to the eigenfrequencies of the modes; the angular speed of the wheelsets correspond to a traveling speed of $v_0 =200$ km/h. The convergence check shows, that the behaviour below 500 Hz can be sufficiently described using 4 bending modes. However, at 600 Hz one can see deviations at the curve for the wheelset 2. The consideration of 11 bending modes provides a sufficient modeling of the vehicle below 1000 Hz; no difference between the curves for 11 and 16 bending modes can be seen. This shows, that for a sufficient modeling of the vehicle in the frequency range below 1000 Hz
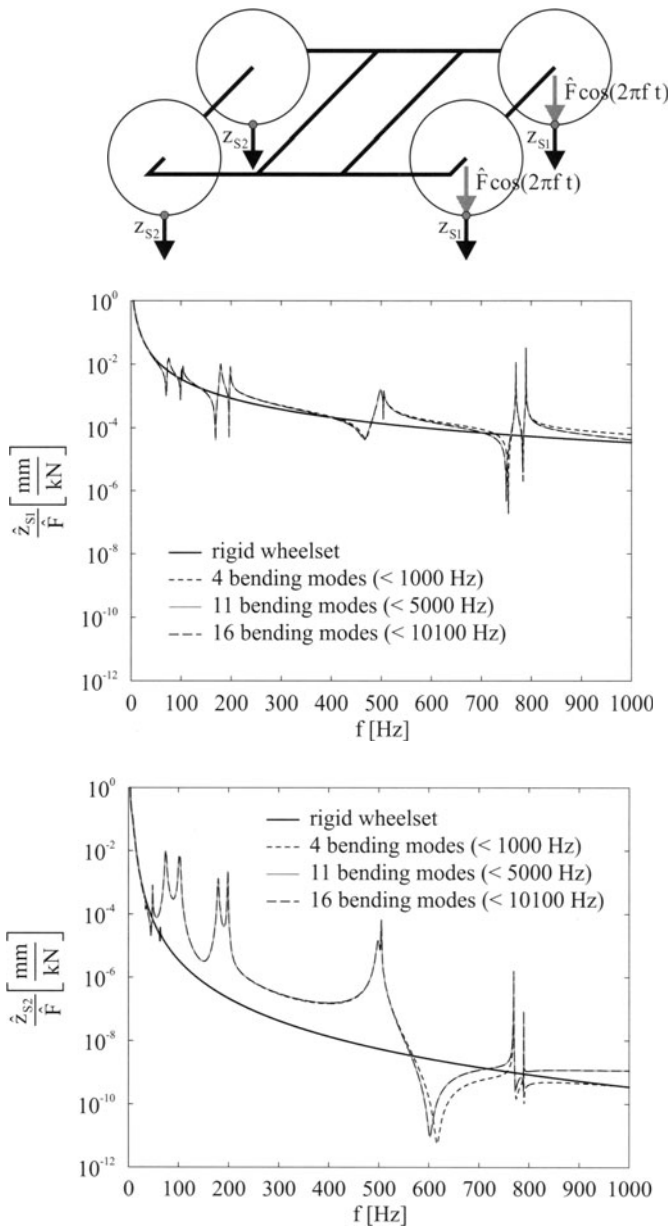
Fig. 6. Convergence check by means of frequency response functions

the eigenmodes of the free wheelset corresponding to the eigenfrequencies located below 5000 Hz have to be taken into account. This comparatively high number results from the fact, that the eigenmodes of the free wheelset have to be used, because no geometrical boundary conditions for the wheelset exist. However, the coupling of the wheelsets and the bogie frame leads to an interaction of the modes.

This analysis also shows the influence of the wheelset elasticity. In the ranges around 100 Hz and 200 Hz, distinct peaks of the frequency response functions occur; these peaks cannot be obtained by the assumption of rigid wheelsets. To demonstrate the influence of the gyroscopic effects, the frequency response is calculated for different angular speeds of the wheelset. Here, an antimetric lateral excitation is chosen; to perform a linear analysis, the yaw damping is assumed to be linear and viscous. The results are shown in Fig.7. Also in
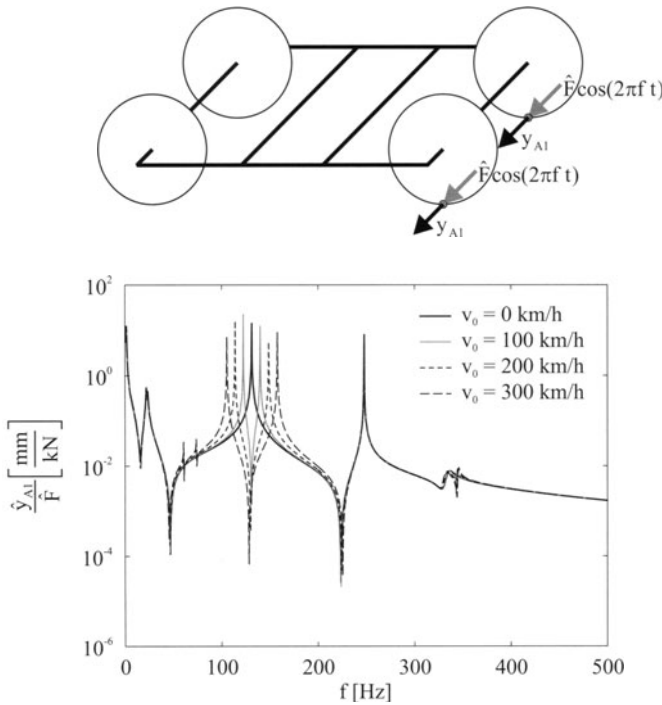
Fig. 7. Influence of the gyroscopic effects shown by frequency response functions

this case, distinct peaks can be seen in the medium-frequency range. With increasing traveling speeds and angular velocities, the peak at 130 Hz, resulting mainly from the bending modes, splits up; this effect is typical for gyroscopic forces. The peak at 220 Hz, however, does not split; it results mainly from the umbrella modes of the wheelset. These modes are rotational symmetric

and do not have a spatial orientation, so the gyroscopic forces have no influence on these motions. The splitted peaks can also be seen in the frequency response functions for the symmetric vertical excitation, see Fig.6.

3.2 Stationary Behaviour of the Vehicle-Track System

For undisturbed rolling of the wheelsets, one obtains the time history of the normal force acting in the wheel-rail contact depicted in Fig.8. The oscil-
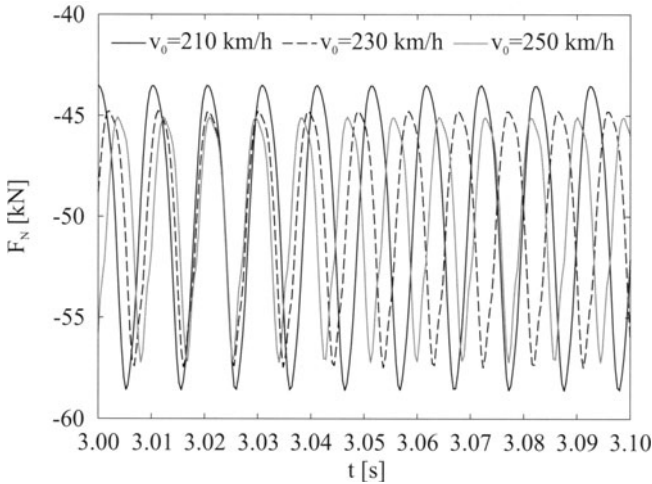
Fig. 8. Normal force for undisturbed rolling

lations of the normal force result from the rolling over the sleepers, which causes a periodic fluctuation of the dynamical properties of the track. It can be seen, that the influence of the sleepers is comparatively weak: The normal forces oscillate in the range between -43.5 kN and -58.5 kN; this means an oscillation between 87% and 117 % of the static wheel load.

As mentioned in the introduction, polygonalized wheels occur due to wear. Fig.9 shows the time history of the normal force for a polygonalized wheel. For this calculation, the right wheel of the wheelset 2 (cp. Fig.2) is assumed to have a harmonic three-periodic polygon with 0.4 mm maximum deviation from the nominal rolling radius. According to [9], even a maximum deviation of the radius of 0.6 mm is tolerable. All other wheels are perfectly circular. Even at the traveling speed of v_0 =210 km/h, lift-offs of the wheel occur. The polygonalized wheel has also an influence on the other, circular wheel of the wheelset. In Fig.10, the normal forces of the wheelset 2 with one polygonalized and one circular wheel are compared; the traveling speed for this calculation is v_0 =230 km/h. It can clearly be seen, that also in the contact of the
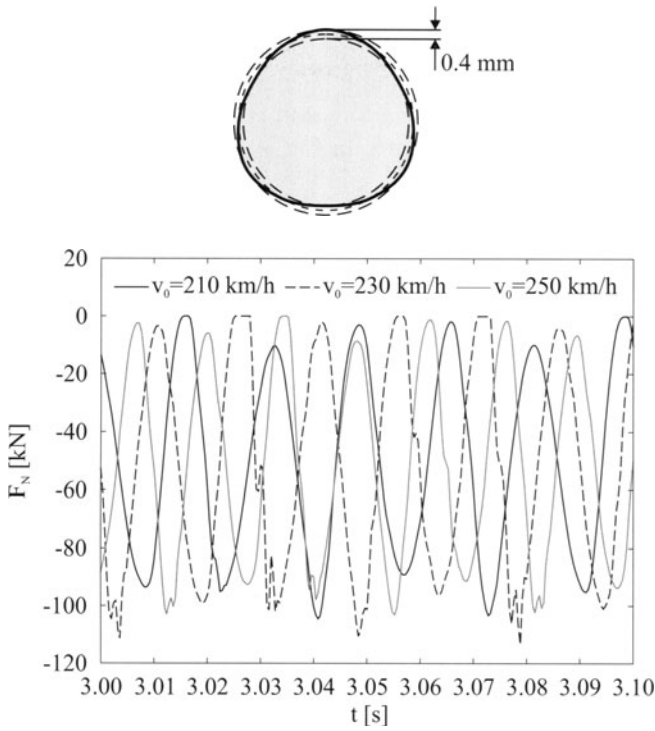
Fig. 9. Normal force for a polygonalized wheel
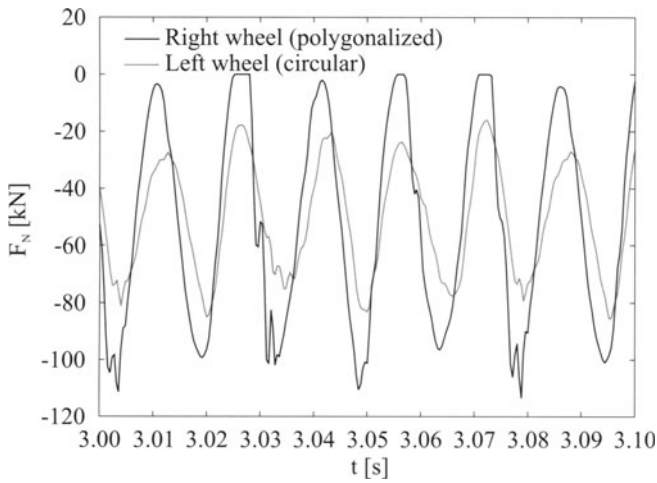
Fig. 10. Normal forces for a wheelset with one polygonalized wheel

circular wheel, large fluctuations of the normal force occur. In this case, the normal force at the circular wheel fluctuates in the range between 60% and 170% of the static wheel load. However, the influence on the other wheelset in the same bogie is very weak; the normal forces in the contact of the other wheelset hardly differ from the forces shown in Fig.8.

3.3   Limit Cycle Behaviour of the Vehicle-Track System

Beyond a certain traveling speed, the so-called critical speed $v_{0,crit}$, the periodic lateral motions do not decrease, but a limit cycle behaviour occurs because of the nonlinearity of the wheel-rail geometry. Fig.11 shows the maximum amplitudes of the lateral motions of the wheelsets 1 and 2 assuming elastic and rigid wheelsets versus the traveling speed. It can clearly be seen,
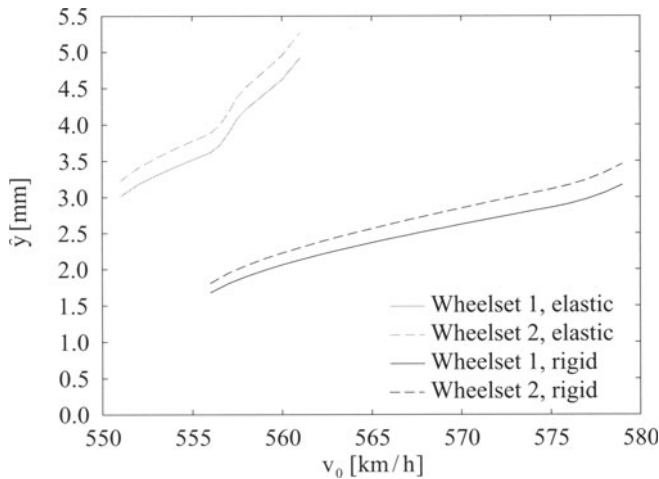
Fig. 11. Limit cycle amplitudes

that the limit cycle behaviour of the elastic wheelsets starts at a lower traveling speed than for the rigid wheelsets. Moreover, the elastic wheelsets perform larger amplitudes than the rigid wheelsets. At the rear wheelset 2, slightly larger amplitudes occur than at the front wheelset 1. This has been known from earlier investigations, see [8].
To explain this behaviour, the influence of the different types of elasticity is investigated. The limit cycles are calculated for pure torsional deformation and for flexural deformation, which includes the bending modes and the umbrella modes. This distinction between torsional and flexural deformations is obvious, because the torsion does not change the shape of the wheelset and has thus no influence on the contact geometry. The resulting limit cycle
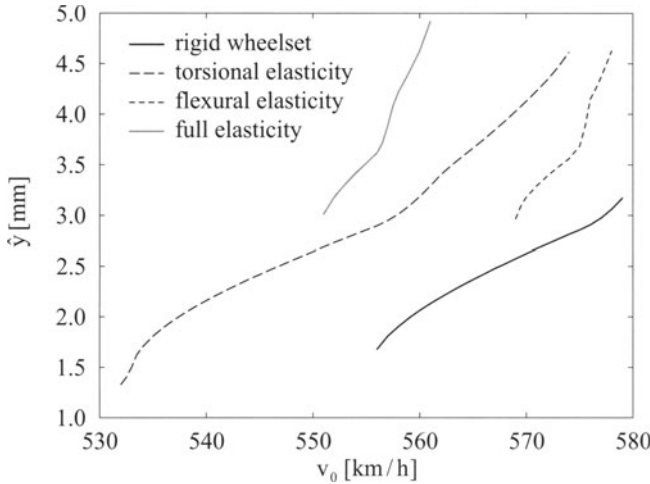
Fig. 12. Influence of different types of elasticity on the limit cycle amplitudes

amplitudes of the lateral motions of the wheelset 1 are shown in Fig.12. It can clearly be seen, that torsion has a very destabilizing influence on the system. The limit cycle behaviour of the torsionally elastic wheelset starts at a critical speed of $v_{0,crit} =532$ km/h, while the rigid wheelset shows this behaviour beyond $v_{0,crit} =556$ km/h.

The flexural elasticity of the bending modes and umbrella modes shifts the critical speed up to $v_0 =569$ km/h. However, the amplitudes of the wheelset are distinctly larger. This behaviour can be explained by the contact geometry. It is known, that the combination of the profiles S 1002 and UIC 60 is very sensitive to even small lateral and rotational displacements. Deformations of the wheelset have an influence on the motion of the wheel rim and thereby on the contact geometry.

For the combination of both types of elasticity, the limit cycle behaviour starts at $v_{0,crit} =551$ km/h. The occurrance of stable limit cycles at lower speeds compared to the rigid wheelset is caused by torsion. However, the critical speed is higher than for pure torsion; this can be seen by the stabilizing influence of the flexural motions. Another effect caused by flexural motions are the distinctly larger amplitudes.

It should be pointed out, that limit cycles for the elastic wheelsets could be found only using a nonlinear theory of the tangential forces; the application of the linear theory of Kalker failed, but not for the rigid wheelsets. Generally spoken, the linear theory yields too large tangential forces, which results in too large deformations of the elastic wheelsets.

3.4   Transient Behaviour of the Vehicle-Track System

To investigate the transient behaviour of the vehicle, the front wheelset 1 is excited by a lateral impact. Fig.13 shows the phase diagrams of the lateral motions of the wheelsets 1 and 2 assuming again elastic and rigid wheelsets. The traveling speed is v_0 =250 km/h, the initial lateral velocity of the front wheelset is $\dot{y}_{WS1}$ =0.5 m/s. As already seen in the investigation of the limit
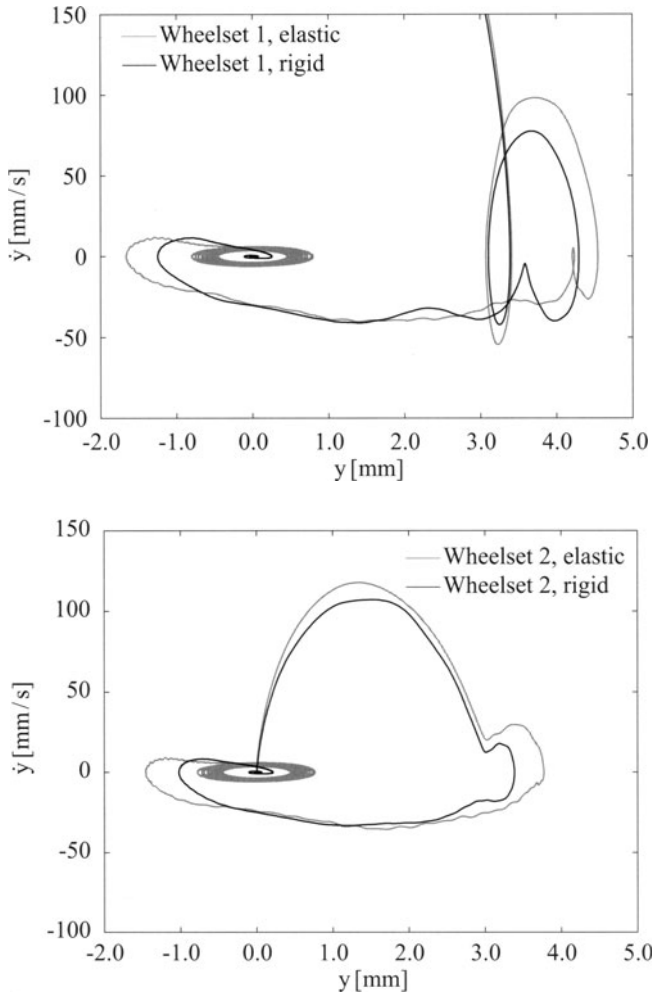
Fig. 13. Influence of the wheelset elasticity on the transient behaviour

cycle behaviour, the elastic wheelsets perform larger amplitudes than the rigid wheelsets. It is remarkable, that the amplitudes of the rigid wheelsets

decrease within 8 s, while the elastic wheelsets show distinct periodic motions with an amplitude of 0.5 mm even after 20 s.

To investigate the influence of the car body, a calculation is performed, where the antimetric motions of the car body (lateral, yaw, and roll motion y, ψ, and φ) are blocked. Fig.14 shows a comparison of the lateral motions of the wheelset 1 assuming a fixed and a free car body; in both cases, the wheelsets are elastic. For large motions, there are only very small deviations between
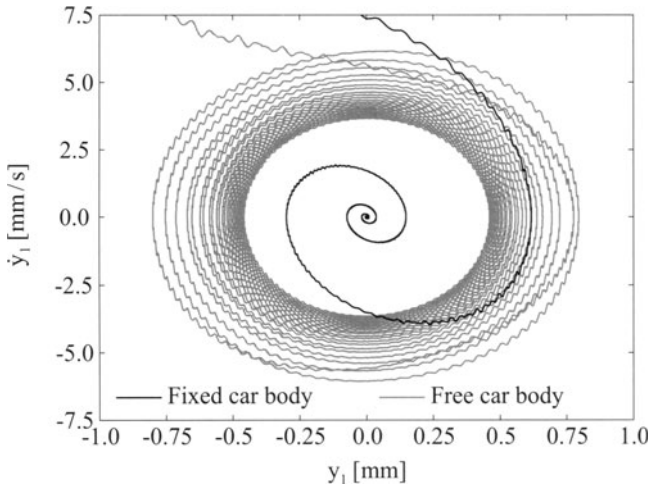
Fig. 14. Influence of the car body on the transient behaviour

the curves. However, in case of small motions, Fig.14 shows, that for a fixed car body the lateral motion of the wheelset diminishes very fast, while the decay of the motions for a free car body is distinctly slower. It should also be pointed out, that the phase curve for a fixed car body represents a time of 4 seconds, while the phase curve for a free car body represents 20 seconds. This behaviour can be explained by the remaining low frequent vibrations of the car body. The frequency of the motions is 1.45 Hz, which corresponds to the eigenfrequencies of the lateral and yaw motions of the car body. The frequency of the superimposed fast oscillations is 115 Hz; this is the frequency caused by passing the sleepers.

Similar as in the last section, the influence of different types of elasticity is investigated. Again, the calculation is performed for pure torsional deformations and for pure flexural deformations; in any case, the car body is free. The results are shown in Fig.15; the curves for the torsional elastic, flexural elastic, and full elastic wheelsets represent a time of 20 seconds, the curve for the rigid wheelset a time of 8 seconds. It can clearly be seen, that the slow decay of the amplitudes is mainly caused by the flexural deformations. The curves
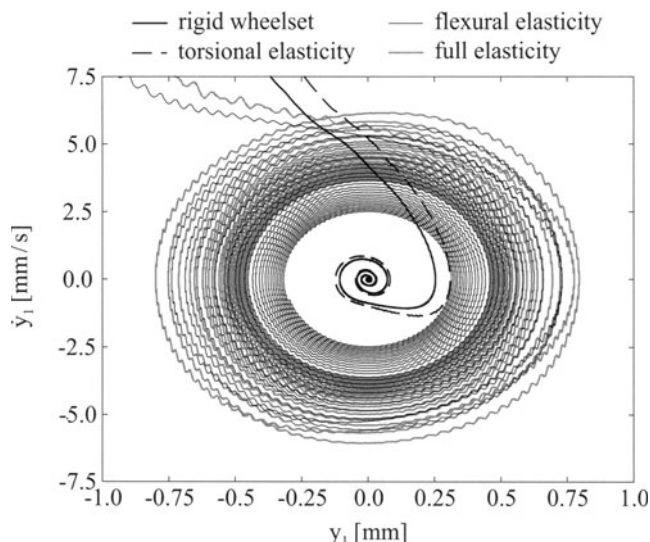
Fig. 15. Influence of different types of elasticity on the transient behaviour

for the rigid wheelset and for the torsionally elastic wheelset hardly differ from each other. The combination of the flexural and torsional deformations leads to larger amplitudes and slower decay than for pure flexural elasticity. The reason of this behaviour can be explained by the influence of the flexural motions on the contact geometry, which is very sensitive, as said before. The flexural deformations are enforced by the lateral and yaw motions of the car body, which has a very large inertia.

It can be said, that the combination of the elasticity of the wheelsets and the motions of the car body cause a very slow decay of the lateral motion of the wheelsets. Due to the fact, that lateral motions cause creepages in the contact, this could lead to increased wear.

4   Conclusion

In this paper, the modeling and the simulation of a railway vehicle with elastic wheelsets has been considered. It is possible to describe the vehicle with rotating elastic wheelsets by linear differential equation with constant coefficients. This simple mathematical structure provides a very easy analysis, which can be used for fast convergence checks of the modal synthesis or the investigation of parameter influences. Furthermore, the chosen description for the rotating elastic wheelsets provides a very simple coupling to the wheel-rail contact module.

The investigation of the frequency response of the vehicle shows distinct peaks in the mid-frequency range, around 100 Hz and 200 Hz for symmetric vertical excitation and around 130 Hz and at 220 Hz for antimetric lateral

excitation. These peaks result from the wheelset elasticity, they cannot be obtained using an MBS consisting of rigid bodies. The frequency response functions calculated with varied numbers of modes can be used for a convergence check. This check shows, that the consideration of the modes of the free wheelset, which correspond to eigenfrequencies below 5000 Hz, is sufficient for the simulation of the vehicle motion in the range below 1000 Hz. Since there are no geometric boundary conditions for the rolling wheelset, the eigenmodes of the free wheelset have to be used. The comparatively high number of necessary modes results from the coupling between the wheelsets and the bogie frame. The increasing angular speed of the wheelsets leads to a splitting of the resonance peaks caused by the bending modes, which is typical for gyroscopic effects.

Furthermore, the interaction of the vehicle with elastic wheelsets and an elastic track has been investigated. The influence of the sleepers is comparatively weak; it leads to fluctuations of the normal force in the range between 87% and 117% of the static wheel load. Polygonalized wheels, however, cause extremely strong fluctuations of the wheel loads. At a traveling speed of $v_0 =210$ km/h, lift-offs and peaks of the normal force with more than 200% of the static wheel load occur at the polygonalized wheel. A single polygonalized wheel also has an influence on the other wheel of the same wheelset. Even if the second wheel is perfectly circular, the normal force oscillates in the range between 60% and 170% of the static wheel load. However, the second wheelset in the same bogie is not influenced by the polygonalized wheel.

The investigation of the limit cycle behaviour shows, that the limit cycle motion occurs at slightly lower traveling speeds for an elastic wheelset than for a rigid wheelset. However, the elastic wheelsets perform distinctly larger amplitudes of the lateral motion than the rigid one. The influence of different types of elasticity, namely the torsional and the flexural elasticity, on the limit cycle behaviour has been investigated. The torsional motions have a very destabilizing effect on the wheelset; for pure torsional elasticity, the critical speed drops to distinctly lower traveling speeds. The flexural elasticity shifts up the critical speed, but when the limit cycle behaviour starts, distinctly larger amplitudes occur. The combination of both types of elasticity leads, as said before, to a slightly lower critical speed and larger amplitudes.

The investigation of the transient behaviour shows, that the combination of the wheelset elasticity with the motions of the car body has a significant influence on the wheelset motions. If only one of these two effects is taken into account, one obtains a comparatively fast decay of the lateral motions of the wheelset. The combination, however, leads to lateral wheelset amplitudes of 0.5 mm, which decrease extremely slow. A comparison of the influence of the torsional and the flexural elasticity on the transient behaviour shows that the slow decay is mainly caused by the flexural motions.

References

1. Popp, K., Kruse, H., Kaiser, I. (1999) Vehicle-track dynamics in the mid-frequency range. Vehicle System Dynamics, 31(5–6): 423–464
2. Morys, G. B. (1994) Zur Entstehung und Verstärkung von Unrundheiten an Eisenbahnrädern bei hohen Geschwindigkeiten. Dissertation TH Karlsruhe
3. Kruse, H. (2002) Modellgestützte Untersuchung der Gleisdynamik und des Verhaltens von Eisenbahnschotter. Dissertation Universität Hannover, to appear
4. Polach, O. (1999) A Fast Wheel-Rail Forces Calculation Computer Code. Proceedings of the 16th IAVSD Symposium, Pretoria, South Africa
5. Fingberg, U., Popp, K. (1991) Experimentelle und theoretische Untersuchungen zum Schallabstrahlverhalten von Schienenrädern. Deutsche Forschungsgemeinschaft, Abschlußbericht zum Forschungsvorhaben Po 136/5-2
6. Fingberg, U. (1990) Ein Modell für das Kurvenquietschen von Schienenfahrzeugen. Fortschritt-Berichte VDI, Reihe 11, Nr. 140. Düsseldorf, VDI-Verlag
7. Ripke, B. (1988) Lateral- und Vertikalschwingungen einer auf diskreten Schwellen gelagerten Schiene unter harmonisch veränderlicher Einzellast. Bericht 192, Institut für Luft- und Raumfahrt, TU Berlin
8. Moelle, D. (1990) Digitale Grenzzykelrechnung zur Untersuchung der Stabilität von Eisenbahndrehgestellen unter dem Einfluss von Nichtlinearitäten. Dissertation TU Berlin
9. Knothe, K. (2001) Dynamik und Festigkeit - Reflexionen zu Eschede. Ringvorlesung "Theorie und Praxis". TU Berlin, 30.1.2001.

Wavy Wear Pattern
on the Tread of Railway Wheels

Michael Küsel and Eberhard Brommundt

TU Braunschweig, Institut für Dynamik und Schwingungen,
Postfach 3329, 38023 Braunschweig, Germany

Abstract. Railway wheels suffer wear, irregular wear pattern develop on their treads. Travel comfort and safety are reduced, expensive reprofiling becomes necessary. A common explanation for the evolution of the wavy wear pattern is seen in an interaction of fast motions of vehicle-track dynamics and slow wear processes. To study this evolution a model will be developed here. The basic idea: The instantanious abrasion in the contact zone is governed by the actual contact conditions. Starting from small initial irregularities the evolution of wavy wear on the treads is simulated numerically. Depending on the running conditions the wear pattern differ with respect to profile and waviness. The results correspond to observations.

1 Introduction

In railway traffic the contact between wheel and rail has a significant influence on travelling comfort and safety. Therefore, the treads of railway wheels and especially the development of wear pattern on the treads are of great interest to the railway companies and a lot of research has been done on this subject[1].

Among others *Zobory* [23] and *Kim* [4] have investigated the development of the profile of the wheels (in the axial direction). The form of the profile has got a strong influence on the lateral dynamics, running stability and running through bends. Out-of-roundnesses (in the circumferencial direction) as they have been studied by *Meinke* and *Meinke* [13], *Morys* [15], *Meywerk* [4] and *Brommundt* [5] have an influence on the vertical and longitudinal dynamics; the oscillating forces impair comfort and safety.

In this paper the wear on the wheel tread will not be dealt with separately for the profile and the out-of-roundnesses. The wear on the wheel tread in axial and in circumferencial direction will be treated as interrelated. This approach is motivated by observations: The wheels of the ICE-train show wavy wear pattern of three troughs in the middle of the tread and only little wear near the flange and the face (see *Zacher* [11], *Morys* [15]). In the "Gotthard Experiment" (see e.g. *Müller* [17]) wear pattern were observed with waves of the 5th harmonic near the flange and waves of the 20th harmonic near the face of the wheel. Depending on the railway vehicle constructions, the

---

[1] *Zobory* [23] inspects the literature on the wear of the profile. *Johannsson* and *Nielsen* [6] present a literature survey on out-of-round railway wheels.

running speeds and the character of the track (straight or curved) different wear pattern develop. How these wear pattern evolve will be sketched in this article[2].

1.1   The Common View of the Evolution of Wavy Wear Pattern

The shape of the wheel tread is described by the radius (see Fig. 2)

$$R(\beta, y, \tau) = R_P(y) + \varepsilon \Delta R(\beta, y, \tau) \,. \tag{1}$$

Here is R_P the wheel profile of the ideal "new" wheel and ΔR denotes some deviation. ΔR is a function of the surface coordinates (y, β) and changes slowly with time τ due to wear. The non-circular constituents of ΔR cause fast oscillating motions of vehicle and track. Two time scales, the fast time t and the slow time τ, are introduced. The evolution of wavy wear pattern on the tread of railway wheels is conceived as an interaction of fast and slow processes (Fig. 1):
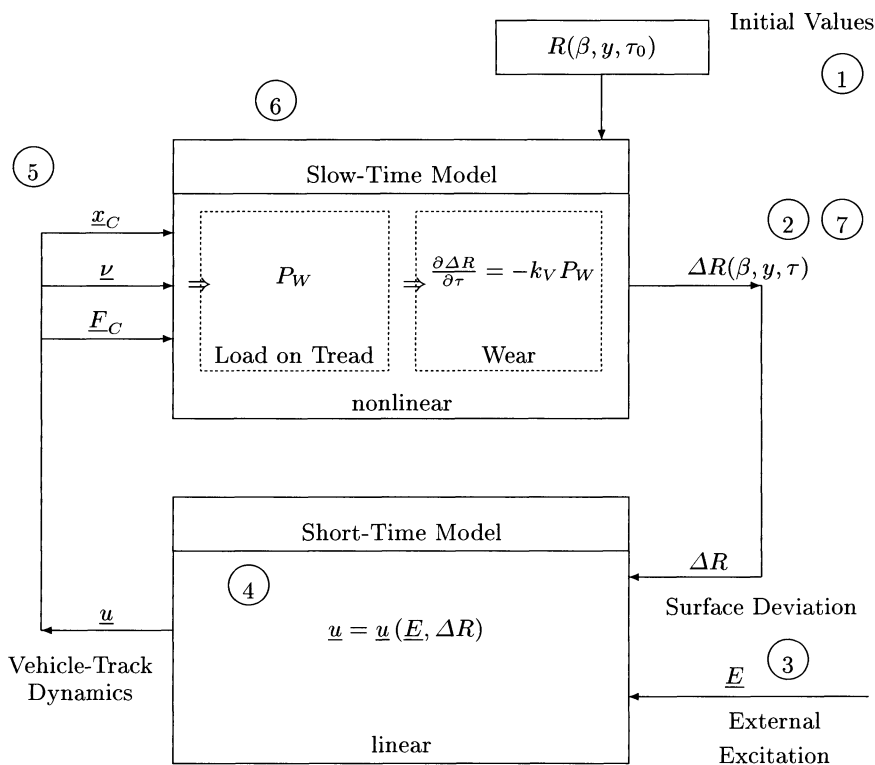
Fig. 1. Interaction of fast- and slow-time dynamics

---

[2] For the full text see Michael Küsel: "Wellige Verschleißmuster auf den Laufflächen von Eisenbahnrädern". To be published.

The initial shape ① of the wheel[3] at $\tau = \tau_0$ contains only a small deviation from the ideal round tread surface ②. The rolling wheels vibrate on the track due to the surface deviations and to external excitations $\underline{E}$ by the motion of the carriage ③. These vibrations of vehicle and track are functions of the fast time t ④. They go along with oscillating contact forces $\underline{F}_C$ and creep $\underline{\nu}$ at the contact point (coordinates $\underline{x}_C$) ⑤. These forces and creep cause wear on the wheel surface ⑥. The wear pattern changes. This closes the feedback loop ⑦. After running distances of approximately 100.000 km the surface deviations exceed a limit and reprofiling becomes necessary[4].

1.2 Shape of Wheel Surface

The surface deviation ΔR will be described in axial direction by a polynomial with respect to the lateral coordinate y and in the circumferencial direction by a Fourier polynomial with respect to the angle β:

$$\Delta R(\beta, y, \tau) = \sum_{n=-N}^{N} \sum_{m=0}^{M} \Delta r_{m\,n}(\tau)\, y^m e^{j\,n\,\beta} \ . \tag{2}$$

The coefficients $\Delta r_{m\,n}(\tau)$ of these sums depend on the slow time τ.
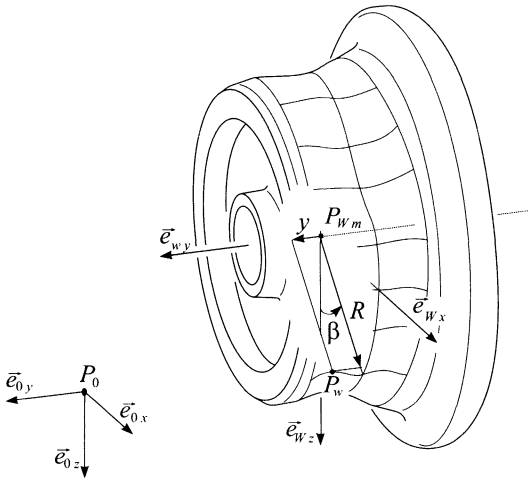
For the radius R given by (1) and (2) the surface of the wheel can be formulated with respect to a fixed reference system $(\underline{e}_0, P_0)$ by the position vector

$$\boldsymbol{r}_W = \underline{r}_{Wm}\underline{e}_0 + [R\,sin(\beta), y, R\,cos(\beta)]\underline{e}_W \ , \tag{3}$$

---

[3] For simplicity we only look at a single wheel.

[4] These running distances depend strongly on the operating conditions, cf. *Müller* [17].

and the surface normal vector reads

$$n_W = \frac{\partial r_W}{\partial \beta} \times \frac{\partial r_W}{\partial y} \, . \tag{4}$$

Similar expressions hold for r_R, n_R, the position vector and the normal vector of the rail.

2   Fast-Time Dynamics

The model for the fast-time dynamics has to describe the motions of the vehicle and the track, and to trace the contact between wheel and rail.

2.1   Vehicle-Track Model

The vehicle-track model for the fast-time dynamics (Fig. 3) consists of a single wheelset, which is running on the track and is attached by spring-damper pairs to the bogie. The bogie is running with constant velocity v_0 and is guided.
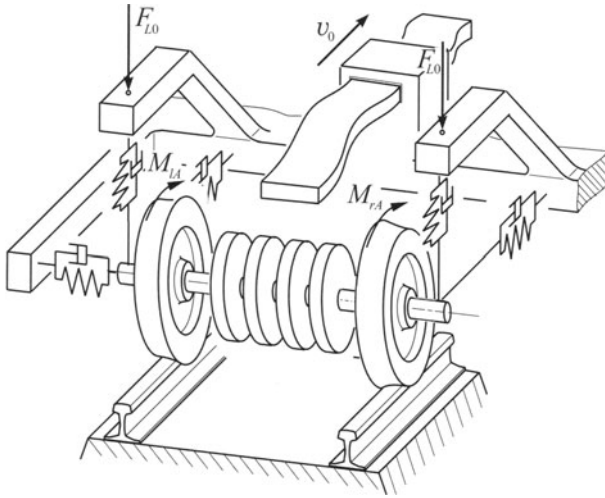
Fig. 3.   Vehicle-track model

The wheelset has 15 DOF (coordinates $\underline{u}_W$), i.e. displacements of the wheel centers, rotations of the wheels, inclinations of the hubs and rims. The deformations of the axle and the wheel disks obey certain shape functions. The equations of motion for the vehicle are established by the Lagrangian formalism.

The track consists of two infinite continuous beams bedded on Winkler foundations. Lateral, vertical and longitudinal displacements plus torsion are taken into account (coordinate matrix $\underline{u}_R$). For harmonic excitation at the point of contact to the wheel, the partial differential equations for the rails are reduced to dynamic stiffnesses.

2.2   Contact Conditions

Vehicle motion and track displacement are coupled by the contact between the wheels and the rails. For the point of contact can be calculated 1. the conditions for the contact geometry, 2. the relative speeds, 3. the contact forces.
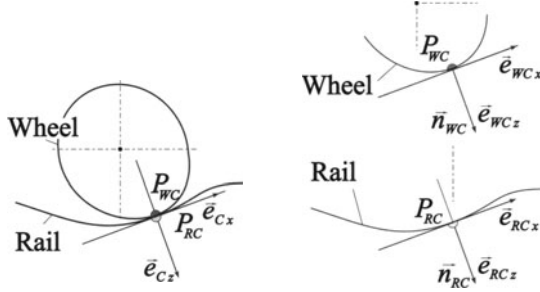
Fig. 4. Geometrical contact conditions

1. As depicted in Fig. 4 wheel and rail touch each other tangentially at one contact point. This leads to the following vector equations:

$$n_{WC} \times n_{RC} = 0 \,, \tag{5}$$

$$r_{WC} - r_{RC} = [0,0,0]\, \underline{e}_C \,. \tag{6}$$

The normal vectors of the wheel and the rail surface have to point into the same direction. The position vectors have to be equivalent. ¿From these conditions one gets for the contact point the contact coordinates x_C and the angular orientations of the contact coordinate systems $\underline{e}_C = \underline{\underline{D}}_C \underline{e}_0$ of wheel and rail.

2. At the contact point hold the relative velocities, the creep ν_x and ν_x,

$$\nu_x = \frac{v_{WC\,x} - v_{RC\,x}}{v_0} \,, \tag{7}$$

$$\nu_y = \frac{v_{WC\,y} - v_{RC\,y}}{v_0} \,, \tag{8}$$

and the spin creep ϕ:

$$\phi := \frac{\omega_{WC\,z} - \omega_{RC\,z}}{v_0} \qquad \text{with} \qquad \underline{\omega} = \underline{\dot{\underline{D}}}_C \underline{\underline{D}}_C^{-1} \tag{9}$$

3. The contact forces are expressed by the linear theory of *Kalker* [9] :

$$\underline{F}_C = \underline{\underline{C}}\,\underline{\nu} \qquad \text{with} \quad \underline{F}_C^T = [F_{C\,x}, F_{C\,y}, F_{C\,z}] \,, \quad \underline{\nu}^T = [\nu_x, \nu_y, \phi] \,. \tag{10}$$

In general, the equations (4 - 9) for the contact are highly nonlinear. By introducing the small parameter ε and the assumption of small surface deviations and displacements the equations are linearized systematically.

2.3   Fast-Time Equations

The equations for vehicle, track and contact are assembled into a single system of equations:

$$\underline{\underline{M}}_2\ddot{\underline{u}} + \underline{\underline{M}}_1\dot{\underline{u}} + \underline{\underline{M}}_0\underline{u} = \underline{f}(\underline{E}, \Delta R)\,, \tag{11}$$

with the state matrix for the vehicle-track dynamics

$$\underline{u}^T = [\underline{u}_W^T, \underline{w}_R^T, \underline{x}_C^T, \underline{\nu}^T, \underline{F}_C^T]\;. \tag{12}$$

The equation describes the forced motion of vehicle and track. The excitations $\underline{f}(\underline{E}, \Delta R)$ on the right hand side contain the influence of the surface deviation (2) and the external excitations:

$$\underline{E}(t) = \sum_m \hat{E}_m e^{j\,m\,s\,\Omega\,t} \qquad (\Omega\ldots\ldots\text{angular velocity})\,. \tag{13}$$

The equations are linearized with respect to the relatively large guided motion of the bogie. Therefore, the fast-time dynamics can be handled in the frequency domain ($\beta \approx \Omega t$). By a Fourier ansatz for the state matrix $\underline{u}$,

$$\underline{u} = \sum_{m,n} \hat{\underline{u}}_{m\,n} e^{j\,(n+m\,s)\Omega\,t}\,, \tag{14}$$

the differential equation (11) reduces to a system of algebraic equations

$$\hat{\underline{u}}_{m,n} = \underline{\underline{H}}_{\Delta R\,m\,n}\underline{f}(\Delta r_{m\,n}) + \underline{\underline{H}}_{E\,m}\hat{\underline{E}}_m\,. \tag{15}$$

The $\underline{\underline{H}}_{(\cdot)}$ are frequency dependend transfer functions. They specify the ratio between the coefficients of the excitations and the forced motions. The main results from the equations are the contact coordinates $\underline{x}_C$, the creep $\underline{\nu}$ and the contact forces $\underline{F}_C$. By these variables the trace of the contact point and the pertinent frictional power on the wheel tread are determined.

Fig. 5 shows an example for the trace of the contact point on the wheel tread and the governing frictional power (The motion does repeat after 7 revolutions of the wheel). This is the essential input to the slow time model.

3   Slow-Time Behaviour

The wavy wear pattern on railway wheels develop slowly with time. As mentioned in the Introduction, the wear will not be equally distributed on the surface of the wheel tread. Therefore the two tasks for the slow-time model are:
- deduce the distribution of the wear load on the wheel surface
- determine for this distribution and for a given wear law the slow evolution of the wheel surface
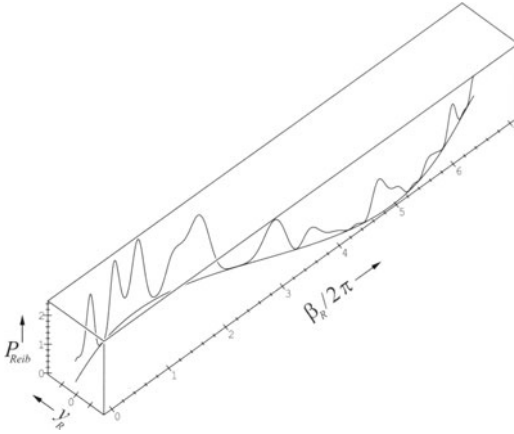
Fig. 5. Trace of the contact point and governing frictional power

The input into the slow-time model is the trace of the contact point on the tread and the governing frictional power as pictured in Fig. 5. The curves are based on the assumption of punctual contact. For the wear model one has to consider an extended contact zone. The contact zone is bordered by the Herzian contact ellipse. Thereby, the curve in Fig. 5 is transfered into a density function. It describes the distribution of the frictional power on the wheel tread. The distribution is assumed - similar to the Herzian normal pressure - to be parabolic over the contact zone (width $2b$).
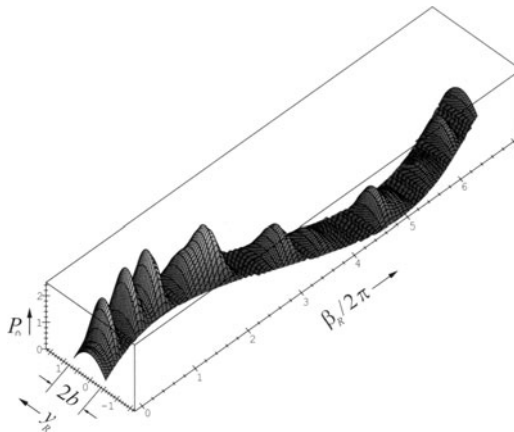
Fig. 6. Distribution of the frictional power

The distribution for the example in Fig. 5 is pictured in Fig. 6. The distribution is approximated by Fourier polynomials in the circumferencial and by Legendre polynomials in the axial direction. Fig. 7 shows the approximation over 7 revolutions and Fig. 8 the final approximation of the frictional power averaged over one revolution.
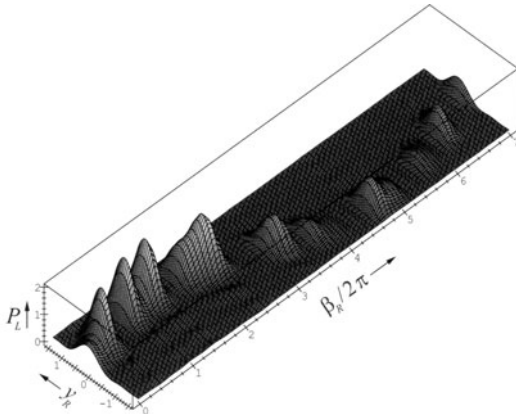
Fig. 7. Approximated frictional power (7 revolutions)
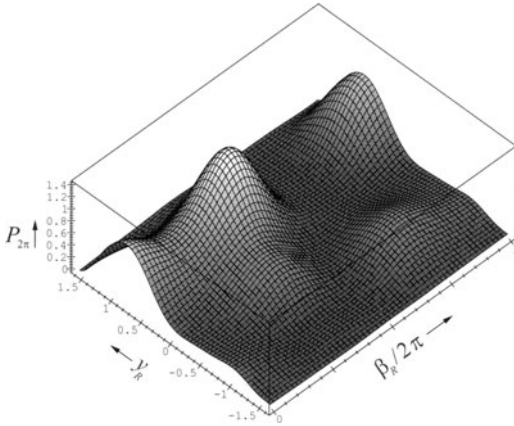
Fig. 8. Approximated frictional power (1 revolution)

This approximation becomes necessary as the frictional power has to be expressed by the same functions as ΔR.

The wear of the wheel surface is assumed to obey:

$$\frac{\partial \Delta R}{\partial \tau} = -k_V P_W \qquad (k_V \dots \text{wear coefficient}). \qquad (16)$$

This wear law is based on the frictional work hypothesis by *Fleischer* [9]: The amount of material removed by wear is proportional to the accumulated frictional energy. Depending on the approximated frictional power the surface deviation develops slowly with time.

4   Evolution of the Coefficients of the Wavy Wear Pattern

Now, the description of each single part of the model has been completed. Nevertheless it is worth looking at the feedback loop of the interaction of

fast and slow processes again (compare Fig. 1 and 9). The feedback loop has been introduced in Section 1.1 in a very general form. Here it is seen that the evolution of the wavy wear pattern can be described by the evolution of the coefficients of the surface deviations. The surface deviations ΔR are expressed by polynomials in y and by Fourier polynomials in β. The coefficients $\Delta r_{m\,n}(\tau)$ - starting from initial values ① - develop slowly with time τ.

Fig. 9. Feedback loop for the evolution of the surface coefficients

The coefficients of the surface deviation $\Delta r_{m\,n}$ and of the external excitations $\underline{E}_m$ are the inputs into the fast-time model ②, ③. By multiplying them with the transfer functions $\underline{\underline{H}}$ one does get the coefficients of the forced motion of the system ④, ⑤. Of special interest are the coefficients of the contact coordinates $\underline{x}_{C\,m\,n}$, the creep $\underline{\nu}_{m\,n}$ and the contact forces $\underline{F}_{C\,m\,n}$ ⑥. These induce the distribution of the frictional power, which will be approximated by Fourier-Legendre polynomials. By numerical integration one does get the coefficients of the surface deviations $\Delta r_{m\,n}$ ⑦. The feedback loop for the evolution of the coefficients of the surface deviation has closed ⑧.

5 Numerical Results

Fig. 10 shows the small initial surface deviations for the left and right wheel. Starting from these initial conditions the bogie is running with constant velocity and on a guided course.
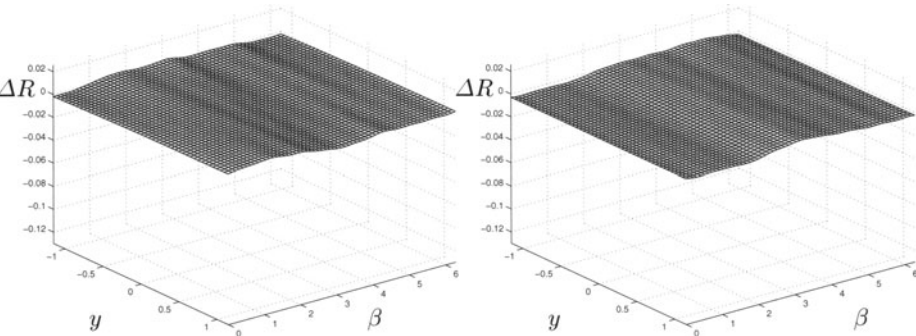
Fig. 10. Initial surface deviations of left and right wheel

Fig. 11 shows schematically the left-right shifted alternating motion of the bogie. The worn out wheel treads in Fig. 12 exhibit distinct wear pattern.
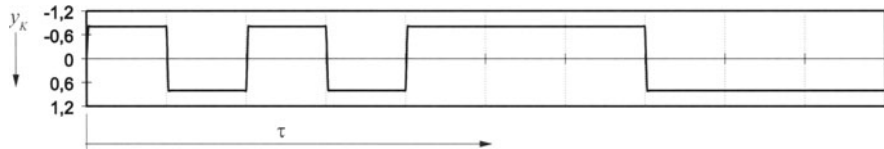
Fig. 11. Example 1: schematic course of the bogie
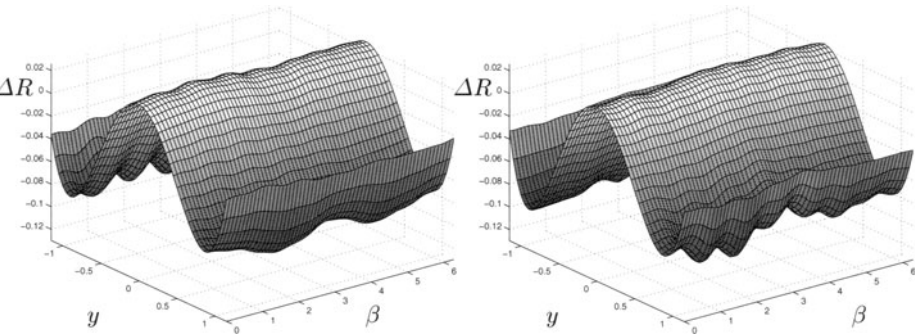
Fig. 12. Example 1: surface deviations at $\tau = 20$

In the regions where the wheel treads had been in contact (left and right side of the tread) deep grooves have formed. These grooves are wavy. The grooves near the wheel flanges show a wavelength of the 2nd harmonic and in the groove near the wheel face of the 7th harmonic.

In the second example the bogie has been guided in a sinusoidal motion (wavelength $\approx 15\pi R_0$) with a alternating small offset.
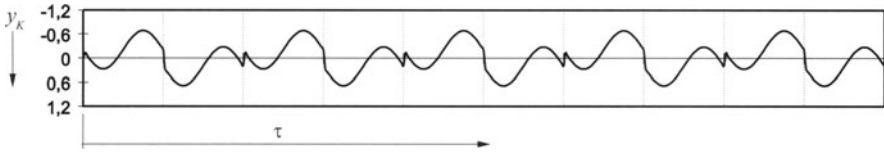
Fig. 13. Example 2: schematic course of the bogie

The worn out wheel treads are shown in Fig. 14.
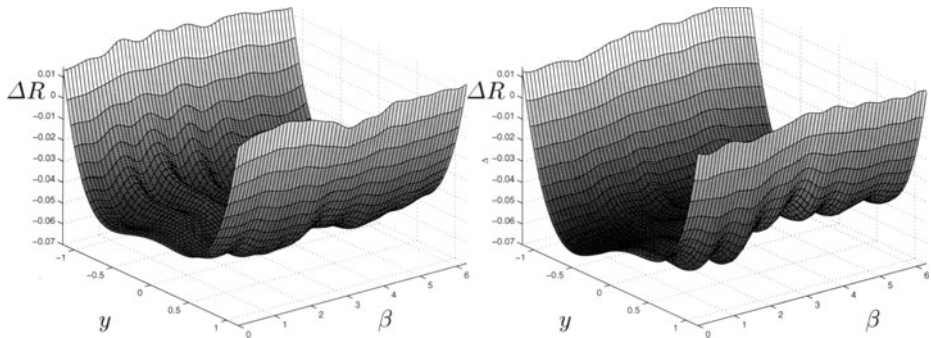
Fig. 14. Example 2: surface deviations at $\tau = 100$

The wear is less severe than in the first example as there has been less constant creep due to the constant offset. Once again the wheel surface is worn out where it has been in contact. In this region the wear is almost uniformly distributed in axial direction. The wheel surface shows, similar to the first example, a wavy wear pattern of the 2nd harmonic near the flange and of the 7th harmonic near the face.

6   Conclusions and Outlook

The wear pattern calculated with this model agree qualitatively with observations. Up to now, a quantitative comparison of the calculated and measured

results has not been attained. The presented model is capable to describe the evolution of wavy wear pattern on railway wheels.

When developing the model it has been paid attention to build up the model by single moduls and to derive the equations starting from very general assumptions. Therefore it is possbile to extend and upgrade the model and to transfer the governing method to related wear problems. Promising future approaches might be a more complex vehicle–track model. Unbalance forces can be included via the external excitations, see equation (13). The guidance of the bogie restricts the motion of the vehicle and the track. To include the influence of the instantanious shape of the tread on the running behaviour is recommendable.

References

1. Meinke, P., Meinke, S. (1999) Polygonization of Wheel Treads Caused by Static and Dynamic Imbalances. J. Sound Vib. **227 (5)**, 979-986
2. Zobory, I. (1997) Prediction of Wheel/Rail Profile Wear. Veh. Syst. Dyn. **28**, 221–259
3. Kalker, J. J. (1990) Three–dimensional Elastic Bodies in Rolling Contact. **2**, Solid Mechanics and its Application. Kluwer Academic Press, Dordrecht, Netherlands
4. Meywerk, M. (1997) Stabilität und Verschleiß bei auf Schienen laufenden Eisenbahnradsätzen. Phd. thesis, Mechanik Zentrum, Technische Universität Braunschweig, Germany
5. Brommundt, E. (1996) Wechselwirkung zwischen Polygonisation und Antrieb bei Eisenbahnrädern. Technische Mechanik **16 (4)**, 273–284
6. Johansson, A., Nielsen, J.(1998) Out-of-Round Railway Wheels – A Literature Survey. **F210**. Department of Solid Mechanics, Chalmers University of Technology, Gothenburg, Sweden
7. Morys, G. B. (1998) Zur Entstehung und Verstärkung von Unrundheiten an Eisenbahnrädern bei hohen GeschwindigkeitenVerschleißgesetz des Rad-Schiene-Systems. Phd. thesis, Universität Karlsruhe (TH), Germany
8. Kim, K.-H. (1996) Verschleißgesetz des Rad-Schiene-Systems. Phd. thesis, Rheinisch–Westfälische Technische Hochschule Aachen, Germany
9. Fleischer, G. (1973) Energetische Methode, zur Bestimmung des Verschleißes. Schmierungstechnik. **4 (9)**, 269–274
10. Müller, R. and Diener, M. (1995) Verschleißerscheinungen an Radlaufflächen von Eisenbahnfahrzeugen. ZEV Glasers Ann. **119 (6)**, 177–192
11. Zacher, M. (1996) Unrunde Räder und Oberbausteifigkeit. ETR. **45 (10)**, 605–610

Rotor Dynamics and Irregular Wear of Elastic Wheelsets

Thomas Meinders and Peter Meinke

Institute B of Mechanics, University of Stuttgart,
Pfaffenwaldring 9, 70550 Stuttgart, Germany

Abstract. High-speed trains in Germany often suffer from vibrations of the car body in the so-called medium-frequency range (30-300 Hz), also known as 100 Hz droning noise. The reason of this unpleasant phenomenon is the development of out-of-round wheels. This paper will discuss the influence of initial out-of-roundness of wheels as well as the influence of wheelset unbalances at high speeds on this wear process. Using a modular approach the model of the wheelset based on the method of flexible multibody systems is coupled to the rail with a complex wheel-rail contact module. In order to account for the long-term wear effects on the wheels the mechanical model is extended by a long-term wear model using a wear-feedback loop in a different time scale. The presented results of wear simulations are subsequently used to discuss the influence of initial out-of-roundness and wheelset unbalances at high speeds on the order and magnitude of developing wear patterns.

1 Motivation

The introduction of the high speed train Intercity Express (ICE) in Germany (1991) has led to new and sometimes only poorly understood problems due to high speed. One of these new problems was easily noticed by passengers due to its loud and disturbing droning noise. The responsible structural vibrations of the car body were excited by out-of-round wheels, which obviously lost their original round shape under the influence of irregular wear. Consequently the changing wheel profiles were causing accelerated wear such that the wheels had to be reprofiled after reaching a critical limit. The characteristic frequency for the excited vibrations of the car body was in the range of 70–100 Hz, which is in the so-called medium frequency range (30–300 Hz).
In order to analyse the rotor dynamics and possible mechanisms for the wear development, an appropriate approach to model flexible bodies in the medium-frequency range has to be selected. Combining the advantages of rigid multibody systems and finite element systems a suitable method is available to account for the first eigenmodes of the wheelset in the questionable frequency range.

2 Flexible Multibody Systems

The method of multibody systems using a minimum set of generalized coordinates has proven to be a very suitable and successful method for the analysis

of constrained mechanical systems, as shown by Schiehlen [18]. In addition to the rigid body approach, where rigid bodies can be connected through massless joints and force elements the extension towards flexible bodies enables the consideration of structural deflections of selected bodies of the multibody system.

The approach to model flexible multibody systems used in this paper is based on the idea assuming large gross motions and superimposed small flexible deformations. For the discretization of the elastic body either local or global shape functions can be used. Due to the flexibility to model even very complex geometric structures, local shape functions (finite element method) have been chosen. The deformation of the structure is described by the modal approach, i.e. through space dependent mode shapes and time dependent modal coordinates, as presented by Kim and Haug [5], Likins [8] and Shabana [19]. The flexible body approach is widely used in vehicle dynamics, see e.g. Claus and Schiehlen [2] and Ambrósio and Gonçalves [1].

2.1   Kinematics and Dynamics

The rigid body motion of a flexible body can be described by introducing a moving reference frame subject to large translational and rotational motions. The position of the reference frame j is then given by the position vector $\mathbf{r}_j(t)$ and the rotation matrix $\mathbf{S}_j(t)$ relative to an inertial frame, see Fig. 1.
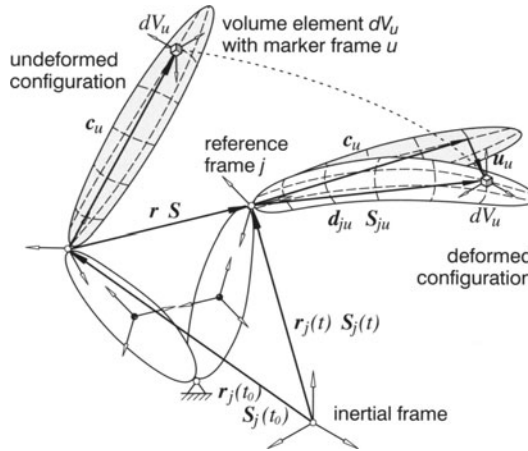
Fig. 1. Definition of deformation and reference vector

The elastic deformations of the body, considered to be small, are described by a displacement vector $\mathbf{d}_{ju}$ and a rotation matrix $\mathbf{S}_{ju}$ linearized with respect to the reference frame. The position of a volume element dV_u relative to the

reference frame j is given by

$$\mathbf{d}_{ju}(\mathbf{c},t) = \mathbf{c}_u + \mathbf{u}_u(\mathbf{c},t) \ , \tag{1}$$

where $\mathbf{c}_u$ indicates the position of the marker frame u in the undeformed position and $\mathbf{u}_u(\mathbf{c},t)$ for the displacement field.

In order to minimize the number of degrees of freedom of the system, a modal reduction is used. Thus the displacement field $\mathbf{u}$ is described by a linear combination of a time invariant translational mode shape matrix $\mathbf{\Phi}$, containing selected deformation modes, and the vector $\mathbf{q}$ of generalized elastic coordinates, which are functions of time only:

$$\mathbf{u}_u(\mathbf{c},t) = \mathbf{\Phi}(\mathbf{c})\,\mathbf{q}(t) \ . \tag{2}$$

Assuming only small deformations, the rotation matrix $\mathbf{S}_{ju}$, accounting for the orientation of the marker frame u relative to the reference frame j, is accordingly described as

$$\mathbf{S}_{ju}(\mathbf{c},t) = \mathbf{I} + \tilde{\vartheta}(\mathbf{c},t) \ , \tag{3}$$

where $\mathbf{I}$ denotes a [3x3] identity matrix and $\tilde{\vartheta}$ a skew-symmetric matrix due to the rotational elastic deformations. The matrix $\tilde{\vartheta}$ is derived from the vector of small rotations $\vartheta = [\alpha\ \beta\ \gamma]^T$. According to (2) this rotation vector can be expressed as a linear combination of a time-invariant mode shape matrix $\mathbf{\Psi}$ and the vector $\mathbf{q}$ of generalized elastic coordinates

$$\vartheta_u(\mathbf{c},t) = \mathbf{\Psi}(\mathbf{c})\,\mathbf{q}(t) \ . \tag{4}$$

As mentioned before the discretization of the elastic body can be accomplished either by using local or global shape functions. The advantage of local shape functions is, that even complex geometric structures can be described. Thus using the finite element method, the translational mode shape matrix (2) can be expressed by

$$\mathbf{\Phi}(\mathbf{c}) = \mathbf{S}^T \mathbf{A}(\mathbf{c})\,\hat{\mathbf{S}}\,\mathbf{B}\,\mathbf{T} \ , \tag{5}$$

where $\mathbf{A}(\mathbf{c})$ is the element shape function matrix, $\mathbf{B}$ is the Boolean matrix describing the assemblage of the finite elements and $\mathbf{T}$ denotes the modal matrix, summarizing the mode shapes of the structure. The matrices $\mathbf{S}$ and $\hat{\mathbf{S}}$ transform the displacements from the element to the reference frame. The rotational mode shape matrix $\mathbf{\Psi}$ is obtained in a completely analogue manner. Before deriving the equations of motions the absolute position and orientation as well as the absolute velocities and accelerations of the marker frame u attached to the volume element dV_u of the elastic body have to be formulated:

$$\mathbf{r}_u = \mathbf{r}_j + \mathbf{d}_{ju} = \mathbf{r}_j + (\mathbf{c}_u + \mathbf{u}_u) \ , \tag{6a}$$
$$\mathbf{S}_u = \mathbf{S}_j \mathbf{S}_{ju} \ . \tag{6b}$$

The expressions for the absolute velocities and accelerations can be derived using relative kinematics, see Melzer [14]. Applying D'Alembert's principle the equations of motion of a flexible multibody system can be written as

$$\mathbf{M}(\mathbf{y}, t)\,\ddot{\mathbf{y}}(t) + \mathbf{k}_c(\mathbf{y}, \dot{\mathbf{y}}, t) + \mathbf{k}_i(\mathbf{y}, \dot{\mathbf{y}}) = \mathbf{q}_f(\mathbf{y}, \dot{\mathbf{y}}, t)\ , \tag{7}$$

with the mass matrix $\mathbf{M}$, the vector of generalized gyroscopic forces $\mathbf{k}_c$, the vector of stiffness and damping forces $\mathbf{k}_i$, and the vector of generalized applied forces $\mathbf{q}_f$. The generalized coordinates of the system are assembled in the vector $\mathbf{y}$, with the vector of the rigid body motion $\mathbf{y}_r$ and the vector of elastic coordinates $\mathbf{q}$ as sub-vectors, such that

$$\mathbf{y}(t) = [\mathbf{y}_r(t)\ \mathbf{q}(t)]^T\ . \tag{8}$$

Various volume integrals have to be evaluated to obtain the mass matrix. Since small deformations are assumed, these volume integrals can be expanded into a Taylor series of elastic coordinates up to first order. The coefficient matrices of the Taylor series, the so-called shape integrals, are calculated by numerical integration. Since the shape integrals are independent of time, they can be computed prior to time integration by pre-processing. A detailed description of this approach can be found in Melzer [14] and Piram [16].

2.2   Procedure of a Dynamic Analysis for a Flexible System

The procedure of the dynamic analysis of a flexible system is presented in Fig. 2. Starting with the definition of the mechanical model of an engineering system, the elastic parts of a multibody system are discretized using a finite element software, e.g. ANSYS [17]. The resulting data for the mass and stiffness matrix as well as the user-selected eigenmodes of the elastic body is used in a preprocessor, e.g. FEMBS [22] in order to compute the shape integrals describing the elastodynamical behaviour. To gain a maximum of software interoperability, these terms are saved in a standardized format (SID). The equations of motion can be computed by a multibody system code, e.g. NEWEUL [7]. Reading the input-file defining the topological structure of the multibody system and the SID-file containing the information about the elastic body, NEWEUL yields mixed symbolic-numerical equations. The simulation of the system can finally be carried out using standard time integration techniques.

3   Modeling of Rotating Wheelsets

The first step in creating a simulation tool to investigate the development of out-of-round wheels is to set up an appropriate mechanical model of the wheelset. Following the procedure exposed in Fig. 2, the essential steps to model the wheelset type BA 14 of the Deutsche Bahn AG, which is the commonly-used wheelset for the German high-speed train ICE 1, are described in the following subsections.
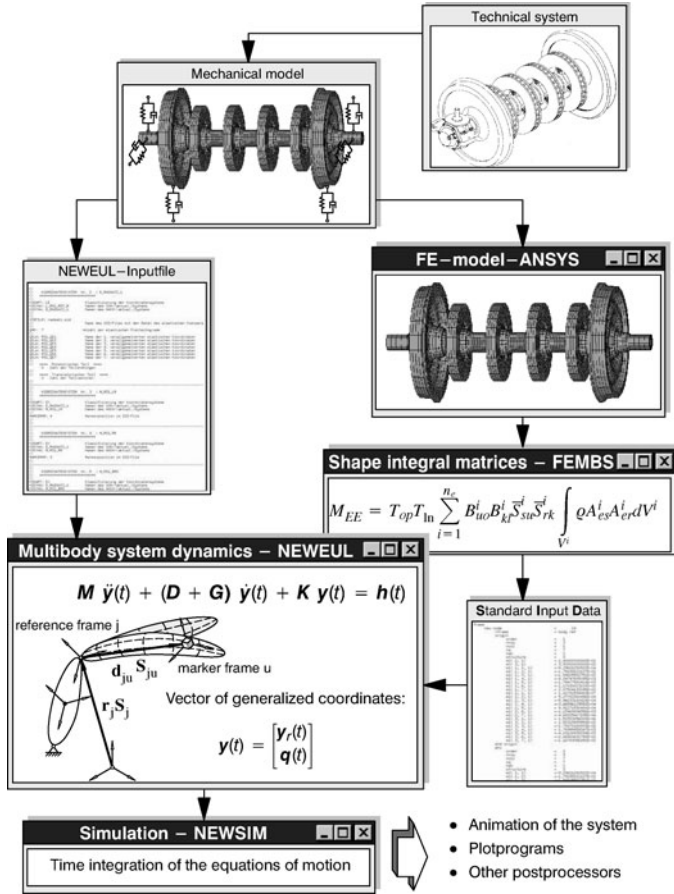
Fig. 2. Procedure of a dynamic analysis of a flexible system

3.1   FE-Model of the Wheelset

The symmetry of the wheelset equipped with altogether 4 disk-brakes is used for the description of the discretized structure by the finite element software ANSYS [17]. In order to gain a maximum of flexibility, the geometric shape of the wheelset is characterized by a set of 54 geometric parameters, as described by Meinders [10]. The 3D finite element structure is generated by rotating a 2D mesh of the wheelset as explained by Meinders [11]. The elements used in this model are SOLID73 from the ANSYS library which provide 6 degrees of freedom for each of the 8 nodes. This is an important requirement for the later use of the finite element data in the flexible multibody system.

Requirements due to the Rotation of the Reference Frame
The connection between the finite element model of the wheelset and the

description of the rigid body model (springs, dampers, bogie coach, rail, etc.) is achieved by a limited number of so-called marker frames u, as shown in Fig. 1. The information about the flexible properties of the body in terms of shape integral matrices is only provided for these selected marker nodes of the finite element model. The reason for this reduction is to keep the size of the overall model and thus the computation time in reasonable limits.

In case of a rotating wheelset the rotation itself is described by a reference frame which is located in the middle of the wheelset on its centerline. Therefore, each node of the structure that is not located on the centerline will rotate with the reference frame with respect to the inertial frame. This can be easily avoided for the interconnecting marker nodes of the primary suspension as well as the marker nodes later needed for the modeling of unbalances by choosing nodes lying directly on the centerline.

The essential wheel-rail contact of the wheelset with its forces and moments is acting on the wheel surface. Therefore, it is not feasible to select one specific node from the surface of the wheels finite element structure since those nodes are rotating relative to the inertial frame.

Realization of the Non-Rotating Wheel-Rail Force

The modeling challenge to realize a non-rotating wheel-rail force acting on the surface of the wheel can be resolved using the following important property of the wheelset: The eigenmode analysis of the FE-structure for the unsupported wheelset (see Fig. 4) as well as for the supported case showed that the wheel-rim has no significant deformation in the frequency range of 0–300 Hz. Thus, the wheel-rim can be treated as a rigid body for the investigation in the medium frequency range.

Using this property of the wheel the degrees of freedom of the wheel-rim elements are constrained and supplemented by four also constrained beam elements such that the motion of the rim is represented by the center-point in the middle of the wheel, as shown in Fig. 3. This center-point is subsequently chosen as a marker node. Thus, the necessary wheel-rail contact forces and moments can act on the wheel-rim even though they are applied to the marker in the middle of the wheel.

Fig. 3. Center-point

3.2   Modal Analysis and Selection of the Elastic Coordinates

The first step in analysing and understanding the dynamical properties of the wheelset is a modal analysis. Subsequently the knowledge about the eigenbehaviour in the medium-frequency range is used to select the eigenmodes needed as elastic coordinates in the flexible multibody system.

The resulting eigenmodes of the unsupported wheelset in the frequency range up to 290 Hz are presented in Fig. 4. At a frequency of 82,5 Hz the first elastic

1st anti–metric torsional mode (82,5 Hz)   2nd symmetric bending mode (188,5 Hz)

1st symmetric bending mode (84,6 Hz)   1st symmetric umbrella mode (235 Hz)

1st anti–metric bending mode (131,8 Hz)   1st symmetric torsional mode (261 Hz)
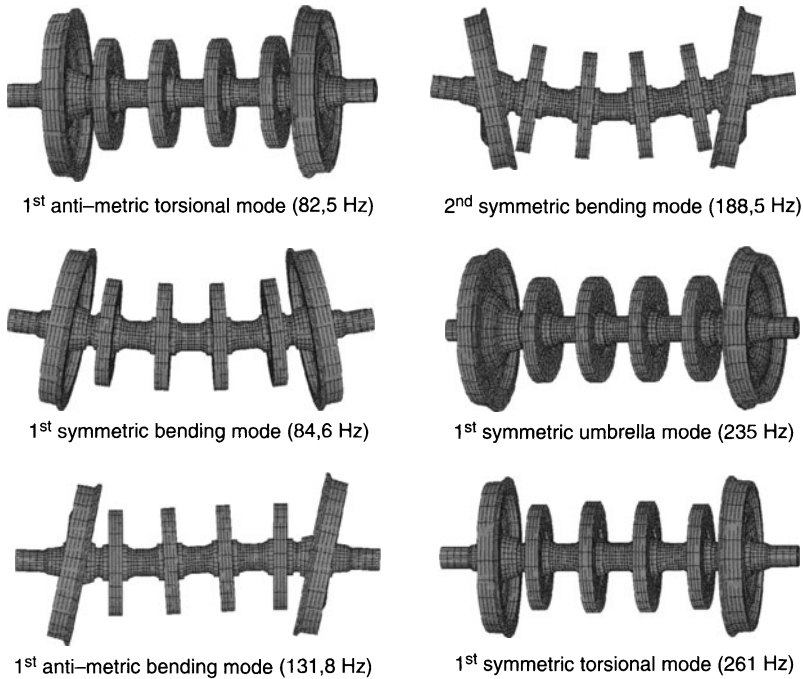
Fig. 4. Eigenmodes of an unsupported wheelset in the frequency range up to 290 Hz

eigenmode of the wheelset is characterized by a torsional motion of one side of the wheelset against the other with a nodal point between the two inner disk-brakes (1st anti-metric torsional mode). The next two eigenfrequencies at 84,6 Hz are the 1st symmetric bending mode in vertical and horizontal direction. At this low frequency wheels and disk-brakes obviously still behave as if they were rigid. This is not true any more for the 1st anti-metric bending mode at 131,8 Hz. At this frequency the flexibility of the wheel membrane influences the movement of the wheels. This can also be found for the 2nd symmetric bending mode at 188,5 Hz, where wheels and disk-brakes bend in opposite directions. At a frequency of 235 Hz the wheel-rims are moving symmetrically along the wheelsets axis. This eigenmode is therefore called 1st symmetric umbrella mode. The 1st symmetric torsional mode at 261 Hz has four nodal points, where the wheels and inner disk-brakes are moving in the same orientation. The last eigenmode in the frequency range up to 300 Hz is the 1st anti-metric umbrella mode at 296,1 Hz.

Based on the knowledge of the eigenbehaviour of the wheelset an accurate selection of type and number of the eigenmodes taken into account for the

inclusions in the flexible multibody system is required. Several simulations with different sets of eigenmodes have shown that the umbrella modes as well as the 1st symmetric torsional mode are not necessary to describe the structural vibrations of the wheelset based on the given boundary-conditions and excitations through unbalances, out-of-round wheels or rail imperfections. As a consequence the following seven eigenmodes are included in the model as generalized elastic coordinates:

- 1st anti-metric torsional mode
- 1st symmetric bending mode (vertical & horizontal)
- 1st anti-metric bending mode (vertical & horizontal)
- 2nd symmetric bending mode (vertical & horizontal)

3.3   Consideration of Static and Dynamic Unbalances

One important parameter for the investigation of the wear development of out-of-round wheels are the unbalances of the wheelset. Thus, the model is prepared for the consideration of static and dynamic unbalances.

Static Unbalances
The magnitude of the static unbalances of the wheels and disk-brakes are based on the regulations of the Deutsche Bahn AG, as described by Meinke and Szolc [13]. According to this the limits for admissible static unbalances

$$U = m\,\varepsilon\,, \tag{9}$$

where m is the mass of the rotating body and ε is the eccentricity (due to manufacturing inaccuracies) with respect to the rotating axes of the rotor are as follows: $U_{wheel} \leq 50\ gm$ and $U_{brake} \leq 16\ gm$.
In the flexible multibody system these static unbalances are modeled with mass-points attached to the corresponding marker frame u of wheels and disk-brakes. The position of the static unbalances corresponds to the mounting regulation by DB AG. As depicted in Fig. 5 the static unbalances of the wheels are mounted opposite to the unbalances of the disk-brakes.

Dynamic Unbalances
Dynamic unbalances can be caused by skew mounted disk-brakes or wheels, an inhomogeneous mass distribution in the rotor of due to measures trying to reduce the size of the static unbalances. Dynamic unbalances are characterized by so-called products of inertia such that the matrix for the mass moment of inertia

$$\mathbf{I} = \begin{bmatrix} I_x & 0 & 0 \\ 0 & I_y & I_{yz} \\ 0 & I_{yz} & I_z \end{bmatrix} \tag{10}$$
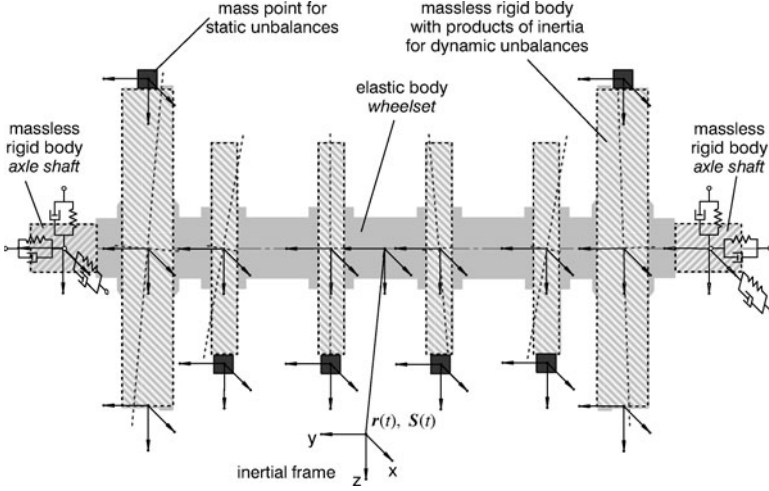
Fig. 5. Consideration of static and dynamic unbalances

contains off-diagonal elements, e.g. the product of inertia I_{yz}. In order to consider this in the flexible model of the wheelset it is important to bear in mind that the principal moments of inertia of the wheels and disk-brakes are already described by the shape integral matrices resulting from the fe-description. Consequently the dynamical unbalances are modeled as massless rigid bodies (see Fig. 5), whose matrix for the mass moment of inertia are missing the principal moments of inertia:

$$\mathbf{I} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & I_{yz} \\ 0 & I_{yz} & 0 \end{bmatrix} \quad . \tag{11}$$

4   Wheel-Rail Contact Module

Railway dynamics are highly influenced by the complex wheel-rail contact situation. Especially for the investigation of wear happening between wheel and rail, a detailed model of this complex contact geometry is essential.

One such detailed model is the wheel-rail contact module of Kik and Steinborn [3], which was originally developed for the use in the multibody system software MEDYNA. Due to its well-defined input-output structure it was possible to extend the multibody system software NEWEUL/NEWSIM [7] with the ability to describe complex railway systems. A detailed report about the integration of the wheel-rail contact module as a force element in the software package NEWEUL/NEWSIM is given by Volle [21].

4.1   Modular Organization of the Contact Module

The principal modular structure of the contact module is shown in Fig. 6. Based on the current position of the wheel j relative to the rail-head i the
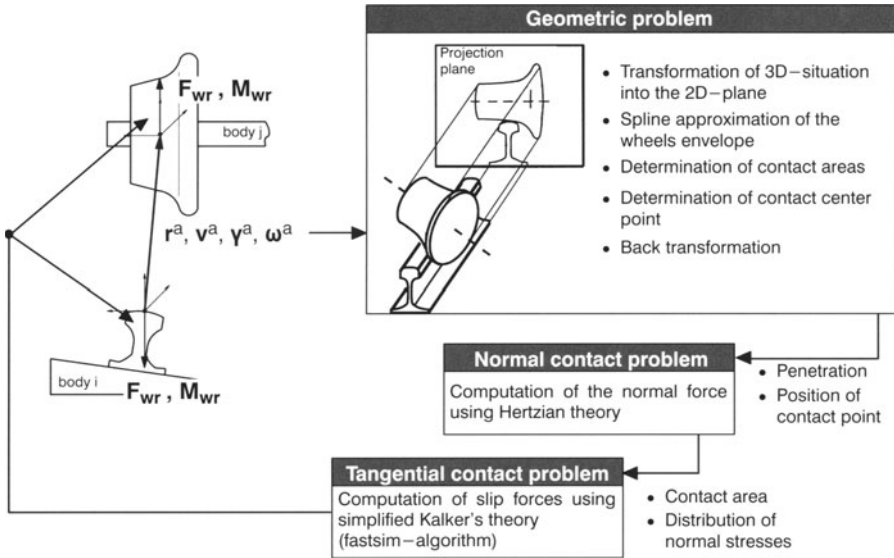
Fig. 6. Modular organization of the contact module

position vector $\mathbf{r^a}$ and the rotation vector $\gamma^{\mathbf{a}}$ as well as the relative velocities $\mathbf{v^a}$ and angular velocities $\omega^{\mathbf{a}}$ serve as the fundamental input parameters for the wheel-rail contact module. Needed for each step of the numerical time integration of the system, the contact module provides the contact forces $\mathbf{F}_{wr}$ and moments $\mathbf{M}_{wr}$ acting between wheel and rail.

As shown in Fig. 6 the contact module is split up into three parts that need to be completed in order to obtain all data for the given contact situation:

Geometric Problem

The contact module enables the use of different spline approximated profiles for wheel and rail. The profiles used for the simulations presented in this paper are UIC60 and S1002. Based on the given relative position of the wheel relative to the rail the 3D contact geometry is transformed into a 2D plane as shown in Fig. 6. Consequently the contact zones and contact points can be determined. Finally the information obtained from this 2D contact situation is transformed back onto the 3D bodies of wheel and rail. The output data obtained from the geometric part of the contact module are essentially the number and position of the contact points, the resulting penetrations in the contact points and the angles of contact.

Normal Contact Problem

The second part of the contact module, the normal contact part, uses the values of the half-axes of the ellipses and the penetration to compute the normal forces based on Hertzian Theory.

Tangential Contact Problem

Finally the tangential contact part of the module computes the tangential forces, twisting moments as well as the slip values. The contact theory used in this part of the model is Kalker's simplified theory, also often referred to as the FASTSIM algorithm.

4.2 Varying Wheel Radii During Time Integration

One important requirement for the use of the wheel-rail contact module is the possibility to describe varying wheel radii depending on the present angular position of the wheel. Since the focus of the wear investigation requires these radii to change over time the radius $r_j(\varphi)$ has yet to be another input value to the contact module.

As depicted in Fig. 7 the local coordinate system for the definition of the
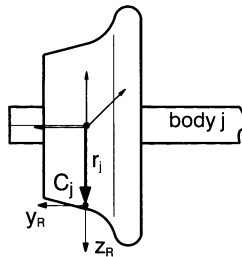
Fig. 7. Definition of wheel profile and wheel radius

wheel profile C_j is not laying in the middle of the wheel, but in the wheels profile itself. By changing the size of the radius $r_j(\varphi)$ the wheels profile in its locally defined coordinate system C_j is changed as a whole.

5   Long-Term Wear Model

The main focus of this paper is the investigation of the wear process of the wheelset. One aspect of this is to determine the amount of mass loss caused by the contact forces and slip values. In order to describe this complex wear process, quite a number of different wear models have been developed, see Kim [4], Specht [20] and Zobory [23]. The wear hypotheses and model for the mass loss used in this paper is presented in Sect. 5.1.

The second part of the wear model is dealing with the long-term effects of the wear. It is therefore necessary to introduce a feedback loop, such that the changing wheel profile is influencing the contact situation between wheel and rail. This influence, often also referred to as long-term behaviour, is happening on a very long time scale that is not accessible through direct time integration.

5.1   Wear Hypothesis and Model for the Mass Loss

The wear model developed for the use together with the contact module from Sect. 4 is based on the following assumptions, see Meinders [12] and Luschnitz [9]:

- The amount of mass loss is proportional to the frictional power (hypothesis of frictional power)
- The wear factor k distinguishes between mild and severe wear
- The frictional power is determined through the contact forces acting in the direction of slip
- Twisting moment and slip are not considered for the calculation of frictional power
- The wear reduces the radius uniformly over the profiles width. It does not change the form of the wheel profile

As mentioned above, the presented wear model is based on the hypothesis, that the loss of material Δm due to wear is basically proportional to the friction work W_R

$$\Delta m = k\, W_R \ . \tag{12}$$

The proportional factor k in (12) is not the same though for all values of frictional power. In fact measurements described by Krause and Poll [6] have shown that the wear factor k is suddenly increasing to a much higher value when a certain frictional power based on the contact area is reached. To reflect this characteristic also shown in Fig. 8, the wear model is distinguishing between mild and severe wear using the following wear parameters:

$$k = \begin{cases} 7 \cdot 10^{-10}\, \frac{kg}{Nm} & : \quad \frac{P_R}{A} \leq 4 \cdot 10^{-6}\, \frac{W}{m^2} \quad \text{mild wear} \\ 2.1 \cdot 10^{-9}\, \frac{kg}{Nm} & : \quad \frac{P_R}{A} > 4 \cdot 10^{-6}\, \frac{W}{m^2} \quad \text{severe wear} \end{cases} . \tag{13}$$

The physical explanation for this sudden increase of the wear parameter k is also given by Krause and Poll [6]: The material surface of the wheels consists of a thin so-called white-itching layer, which shows a higher resistance against wear than the underlying base material. This white-itching layer is transformed out of the base material due to the wear induced impact. As long as there is an equilibrium between the buildup of the white-itching layer and the abrasive effects of wear reducing this layer, this is considered as mild
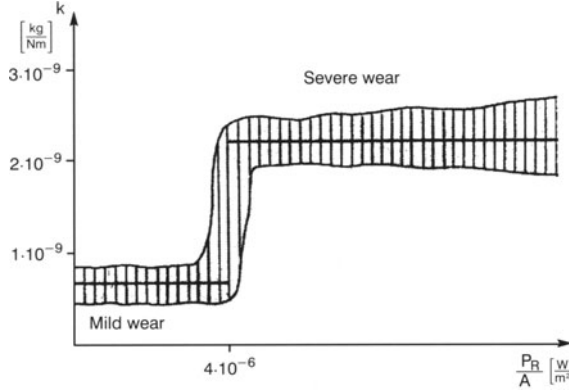
Fig. 8. Wear factor k as a function of frictional power per contact area $\frac{P_R}{A}$ $[\frac{W}{m^2}]$

wear. As soon as the white-itching layer vanishes due to higher frictional power, the much less wear resistant base material is exposed to the wear. This is consequently considered as severe wear.

As mentioned above the frictional power P_R is determined through the contact forces **F** acting in the direction of slip $\boldsymbol{\nu}$, that is

$$P_R = \mathbf{F}\boldsymbol{\nu} \ . \tag{14}$$

In order to obtain the changing wheel radius $\Delta r(\varphi)$, the following relations with ρ as the density of the wheel material and A_i as the size of the contact area during the time interval t_i can be written as:

$$\Delta m = \rho\,\Delta V$$
$$= \rho\,A_i\,\Delta r(\varphi) \tag{15}$$
$$A_i = \Delta U\,b_i \ . \tag{16}$$

Hence, the changing wheel radius $\Delta r(\varphi)$ can be expressed as

$$\Delta r(\varphi) = \frac{\Delta m}{\rho\,b_i\,\Delta U} \ , \tag{17}$$

with b_i and ΔU denoting the width and length of the contact area, as also illustrated in Fig. 9.

Finally, if the wear hypotheses from (12) is used as well as the expression for the length of the contact zone during Δt, that is $\Delta U = V_0\,\Delta t$ with V_0 as the longitudinal speed of the wheelset, the change of the radius for the segment S_{kl} of the discretized wheel surface can be expressed as

$$\Delta r_{kl} = \frac{k\,P_R}{\rho\,b_i\,V_0} \ . \tag{18}$$
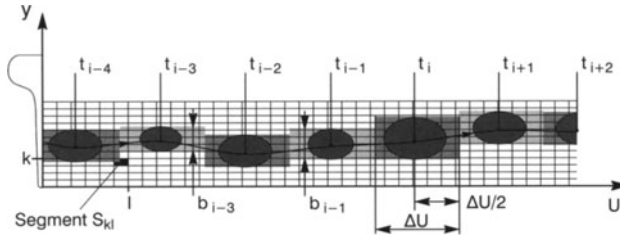
Fig. 9. Distribution of wear over the discretized wheel surface

5.2   Feedback of Worn Profiles in the Sense of Long-Term Wear

The changing profiles of the wheels, that are slowly losing their original shape due to wear, will influence the dynamics of the system. Therefore, it is important to close the feedback-loop as shown in Fig. 10, such that the changing wheel-radius is used by the contact module.
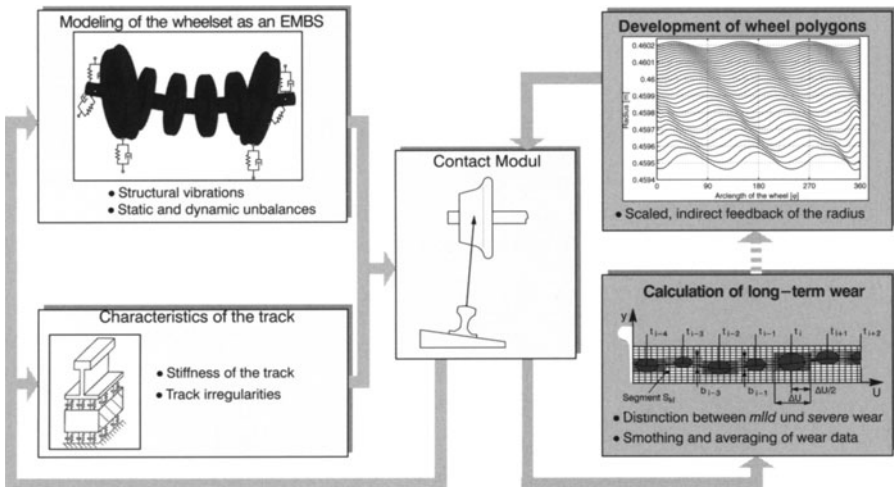
Fig. 10. Feedback-loop of the long-term wear model

In reality the observed phenomenon of developing out-of-round wheels normally takes about $100.000\ km$ in order to show measurable amplitudes of the wheel radius (e.g. $\Delta r = 0,3\ mm$). Even recent computers are not able to simulate the presented model of the wheelset for a corresponding period of time of presumably weeks or months. Thus it is necessary to introduce a time lapse for the occurring wear. The scaling factor $c_w = 10.000$ is enabling wear simulations with relevant amplitudes of the radius, such that the effects of out-of-round wheels on the dynamics of the system can be studied.

The disadvantage of the amplification of the wear is though, that single wear effects gain an undesirable impact. In order to balance the impact of those effects, the changed radii of the wheels are made available to the contact module in terms of an indirect feedback. This is achieved by accumulating the occurring wear for a number of revolutions $\bar{n} = 10$.

6   System and Wear Behaviour of Elastic Wheelsets

Investigating the behaviour of an ICE 1 wheelset equipped with altogether four disk-brakes, the following aspects are of particular interest: Compared to conventional wheelsets of rail vehicles traveling at lower speed the ICE 1 wheelset is exposed to very high rotational speeds. This is particularly interesting since the moments of inertia considerably increased after adding the four disk-brakes to the wheelset axle. Therefore, it is interesting to investigate whether the wheelset running at high speeds should be treated as a rotor, see Sect. 6.1.

The essential question of this paper is the analysis of the long-term wear development due to different excitations. It is well known, see Morys [15], that even new wheelsets do not have a perfectly round shape, but already show some characteristics of out-of-round wheels. Thus the influence of such initial out-of-roundness is analysed in Sect. 6.2.

Another form of excitation possibly effecting the dynamics and consequently the wear development of the wheels are the unbalances already mentioned in Sect. 3.3.

6.1   Eigenbehaviour of the Wheelset

According to the eigenmode analysis in Fig. 11, the bending modes of the
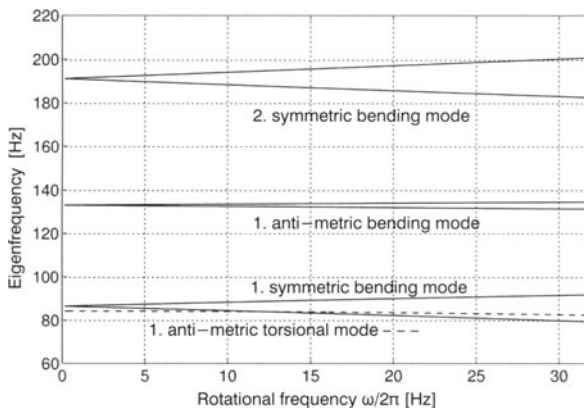
Fig. 11. Bifurcation of the eigenfrequencies over the rotational frequency

rotating flexible wheelset split into concurrent rotating and counter-rotating eigenmodes. This phenomenon also referred to as bifurcation of the eigenfrequencies clearly shows the characteristic behaviour of an elastic rotor. It is also clear from Fig. 11, that the torsional mode is independent from the rotational frequency. The speed corresponding to the rotational frequency range investigated in Fig. 11 is $V = 0 - 300 \, km/h$.

6.2 Wear Development due to Initial Out-Of-Roundness

The simulations of long-term wear showing the influence of the order of initial out-of-roundness have all been carried out with a traveling speed of $V = 256 \, km/h$. The simulation time for all diagrams Fig. 12–15 together with the scaling factor $c_w = 10.000$ (see Sect. 5.2) corresponds to the distance of about 73.600 km. The initial amplitude of the initial out-of-roundness is $\Delta r = 0,02 \, mm$. Beside the excitation through initial out-of-roundness there are no other excitations through unbalances or track irregularities.
The initial out-of-roundness shown in Fig. 12 can be caused by excentric
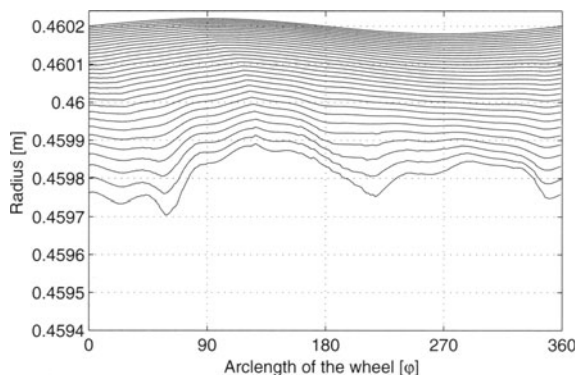
Fig. 12. Influence of initial first order out-of-roundness

mounted wheels. Besides the reduction of the overall radius of about $\Delta r \approx 0,4 \, mm$ the amplitude is increasing up to $\Delta r \approx 0,2 \, mm$. It is clear from Fig. 12, that an initial first order out-of-roundness tends to develop into a second order out-of-roundness.
In contrast to that the second order out-of-roundness as depicted in Fig. 13 is more stable concerning its order. Even though the overall reduction is also about $\Delta r \approx 0,4 \, mm$ this wear development leads to an accelerated increase of the amplitude up to $\Delta r \approx 0,4 \, mm$.
The highest overall wear of all presented diagrams is obviously caused by the initial out-of-roundness of third order, see Fig. 14. The overall loss of the radius is $\Delta r \approx 0,7 \, mm$, whereas the depth of the amplitude is increasing slowly to $\Delta r \approx 0,1 \, mm$.
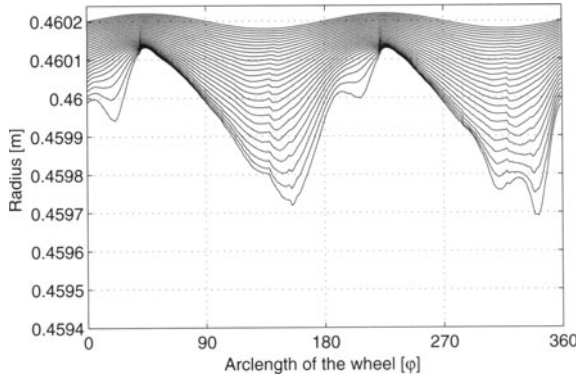
Fig. 13. Influence of initial second order out-of-roundness
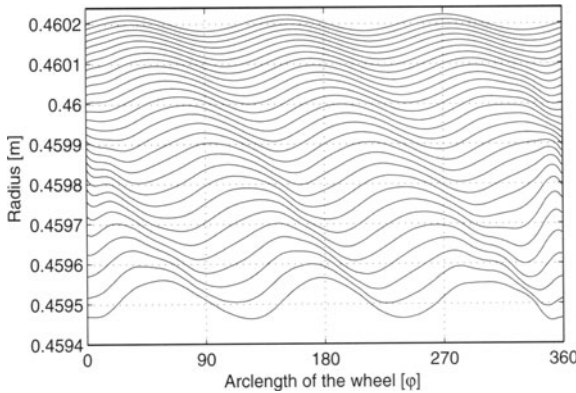
Fig. 14. Influence of initial third order out-of-roundness

The fifth order out-of-roundness in Fig. 15 shows similar characteristic as the second order. While the initial shape in relatively stable in shape, the amplitude of the out-of-roundness is increasing up to $\Delta r \approx 0,2\,mm$.

Summarizing the results of the influence of different initial out-of-roundnesses it can be concluded that higher order out-of-roundness tend to be rather stable concerning their shape whereas the first order out-of-roundness leads to second order out-of-roundness.

6.3   Wear Development due to Unbalances

The initial shape of the wheels for investigating the influence of static and dynamic unbalances is a perfectly round wheel. As before the evolving shape of the wheels as shown in Fig. 16 and 17 does correspond to the distance of 73.600 km with a speed of $V = 264\,km/h$ and without excitation by track irregularities.
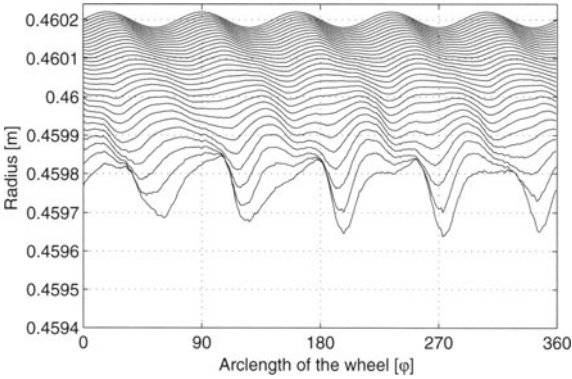
Fig. 15. Influence of initial fifth order out-of-roundness
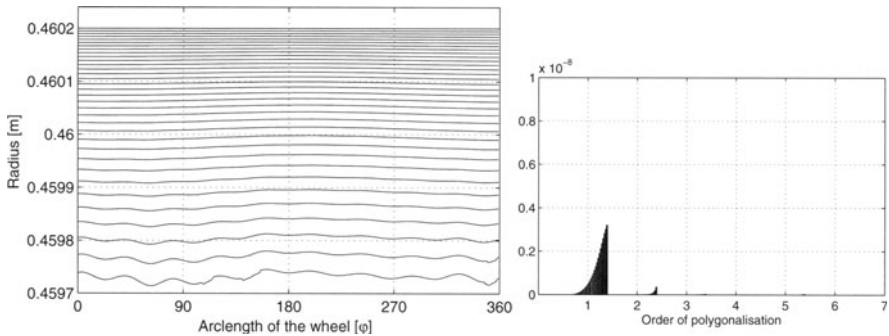
Fig. 16. Wear development due to static unbalances
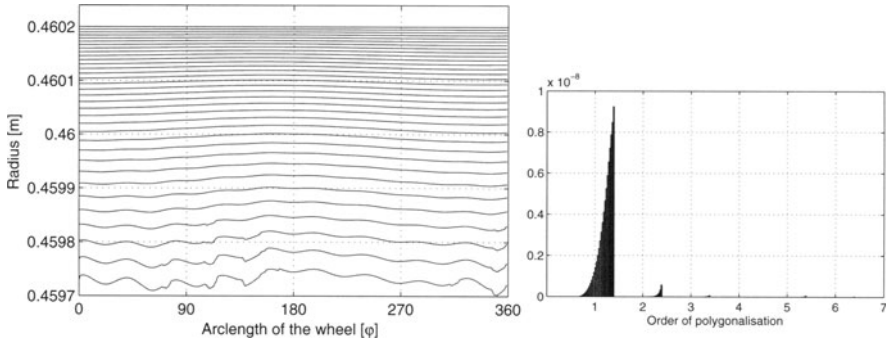
Fig. 17. Wear development due to dynamic unbalances

According to the resulting wear development for the excitation with static unbalances in Fig. 16 as well as for dynamic unbalances in Fig. 17 it can be concluded that unbalances do only have a small influence on the wear development compared to the impact of initial out-of-roundness. For both types of unbalances the first order out-of-roundness is emerging. The diagrams on the right hand side of Fig. 16 and 17 show that dynamic unbalances are leading to a three times faster growth of the first order unroundness than the static unbalances.

7   Summary

The approach of flexible multibody systems is used to analyse the rotor dynamics and the developing irregular wear of elastic railway wheelsets. The modeling challenge of a rotating flexible body exposed to the wheel-rail forces is resolved by applying the contact forces to a center-point, which is coupled to the wheel-rim through constrained equations. A modular wheel-rail contact module is integrated in the multibody system software NEWEUL/NEWSIM. In order to account for the appearing wear of the wheels a long-term wear model with a feedback loop for the changing profiles is presented.

According to the eigenmode analysis which results in seven different eigenmodes of the wheelset in the medium frequency range up to 300 Hz, it is essential to consider the possible deformations of the wheelset. The influence of the rotational speed on the bending eigenfrequencies of the wheelset emphasize the relevance of treating the wheelset as a rotor. Initial out-of-roundness prove to have a major impact on the wear development. The growth of the amplitude of all kinds of out-of-roundness is remarkable. Further, the initial shape of second and higher order out-of-roundness are rather stable concerning their shape. Finally static and dynamic unbalances only have a minor influence on the wear development.

References

1. Ambrósio, J. A.C., Gonçalves, J. P.C. (2001) Flexible multibody systems with applications to vehicle dynamics. Multibody System Dynamics, **6**(2), 163–182
2. Claus, H., Schiehlen, W. (2002) System dynamics of railcars with radial- and lateralelastic wheels. Also published in this volume
3. Kik, W., Steinborn, H. (1982) Quasistationärer Bogenlauf - Mathematisches Rad/Schiene Modell, ILR Mitt. 112. Institut für Luft- und Raumfahrt der Technischen Universität Berlin, Berlin
4. Kim, K. (1996) Verschleißgesetz des Rad-Schiene-Systems. Dissertation, RWTH Aachen, Fakultät für Maschinenwesen

5. Kim, S. S., Haug, E. J. (1990) Selection of deformation modes for flexible multibody dynamics. Mech. Struct. & Mach. **18**(4), 565–586

6. Krause, H., Poll, G. (1984) Verschleiß bei gleitender und wälzender Relativbewegung. Tribologie und Schmierungstechnik 31, RWTH Aachen

7. Kreuzer, E. (1979) Symbolische Berechnung der Bewegungsgleichungen von Mehrkörpersystemen, Fortschrittsberichte VDI-Zeitschrift, Reihe 11, 32. VDI-Verlag, Düsseldorf

8. Likins, P.W. (1967) Modal method for analysis of free rotations of spacecraft. AIAA Journal **5**(7), 1304–1308

9. Luschnitz, S. (1999) Ein Modell zur Berechnung des Langzeitverschleißes bei ICE-Radsätzen, Studienarbeit STUD-176 (Meinders, Schiehlen, Meinke). Institute B of Mechanics, University of Stuttgart

10. Meinders, T. (1997) Rotordynamik eines elastisches Radsatzes, Zwischenbericht ZB-94. Institute B of Mechanics, University of Stuttgart

11. Meinders, T. (1998) Modeling of a railway wheelset as a rotating elastic multibody system, Machine Dynamics Problems **20**, 209–219

12. Meinders, T. (1999) Einfluß des Rad-Schiene-Kontakts auf Dynamik und Verschleiß eines Radsatzes, Zwischenbericht ZB-116. Institute B of Mechanics, University of Stuttgart

13. Meinke, P., Meinke, S., Szolc, T. (1995) On dynamics of rotating wheelset/rail systems in the medium frequency range. Dynamical problems in mechanical systems IV, IPPT PAN, Warsaw, 233–244

14. Melzer, F. (1994) Symbolisch-numerische Modellierung elastischer Mehrkörpersysteme mit Anwendung auf rechnerische Lebensdauervorhersagen, Fortschrittsberichte VDI-Zeitschrift, Reihe 20, 139. VDI-Verlag, Düsseldorf

15. Morys, B. (1998) Zur Entstehung und Verstärkung von Unrundheiten an Eisenbahnrädern bei hohen Geschwindigkeiten. Dissertation, Fakultät für Maschinenbau, Universität Karlsruhe

16. Piram, U. (2001) Zur Optimierung elastischer Mehrkörpersysteme, Fortschrittsberichte VDI-Zeitschrift, Reihe 11, 298. VDI-Verlag, Düsseldorf

17. N.N. (2000) ANSYS User's Manual. Ansys Inc., Houston, Pennsylvania

18. Schiehlen, W. (1997) Multibody system dynamics: Roots and perspectives. Multibody System Dynamics **1**(2), 149–188

19. Shabana, A. A. (1989) Dynamics of multibody systems. Wiley, New York

20. Specht, W. (1985) Beitrag zur rechnerischen Bestimmung des Rad- und Schieneverschleißes durch Güterwagendrehgestelle. Dissertation, RWTH Aachen, Fakultät für Maschinenwesen

21. Volle, A. (1997) Integration eines Rad-Schiene-Kontaktmoduls in die Simulationsumgebung NEWSIM, Diplomarbeit DIPL-67 (Meinders, Schiehlen, Meinke). Institute B of Mechanics, University of Stuttgart

22. Wallrapp, O., Eichberger, A. (2000) FEMBS, An interface between FEM codes and MBS codes. DLR. Oberpfaffenhofen

23. Zobory, I. (1997) Prediction of wheel/rail profile wear. Vehicle System Dynamics, **28**, 221–259

Part III:

Contact, Friction, Wear

On the Numerical Analysis of the Wheel-Rail System in Rolling Contact

Sabine Damme[1], Udo Nackenhorst[2], Anja Wetzel[1], Bernd W. Zastrau[1]

[1] Lehrstuhl für Mechanik, Technische Universität Dresden
[2] Institut für Baumechanik und Numerische Mechanik, Universität Hannover

Abstract. The evaluation of the contact patch and the distribution of stress and strain in rolling contact requires a geometrically and materially nonlinear approach. The presented paper gives an overview over the theoretical background of the Arbitrary LAGRANGian-EULERian formulation of the stationary dynamics of rolling bodies in elastic-elastic contact. Contact formulations for dynamic contact conditions are developed using a velocity oriented approach. The transcription into the Finite Element Method, which is temporarily limited to an elastic material law, is given. For future developments some conceptual comments for the extension for transient rolling contact are made. Numerical examples for three dimensional simulations including rolling over worn profiles demonstrate the applicability of the developed program package. A prospect of started key issues of research and development finalizes the paper.

1   Introduction

Although railway systems are a transportation system with a long tradition still a large number of unsolved problems exists. One of these problems is the assessment of the stress and strain status of the rail and the wheel under rolling contact conditions. Since the mechanical problem is of high complexity because of the contact condition and the material behavior, only a numerical approach seems to be appropriate. In this publication a method is described which has the advantage to enable the engineer to treat both the contact problem and material nonlinearity within one simulation tool. For the judgement of the applicability of other approaches such as the utilization of the Boundary Element Method as proposed e.g. by Kalker [3] and others it shall be referred to an overview article by Knothe et al. [5]. The herewith developed theoretical approach and numerical realization will be confined to elastic material, although plastic deformation of the rough surface has been exploited for an appropriate modelling of the normal contact condition. Neither material deterioration nor plastic deformation or shakedown analysis will be treated in this paper, although the obtainable stress distribution could be exploited for the evaluation of the aforementioned effects, heat generation and wear.
The numerical simulation of rolling contact problems goes back to the 1980s, when computers and computer codes for nonlinear problems started to develop rapidly. The computation of rolling contact phenomena started with

adaptively refined mesh

(a) discretization

(b) tangential surface tractions
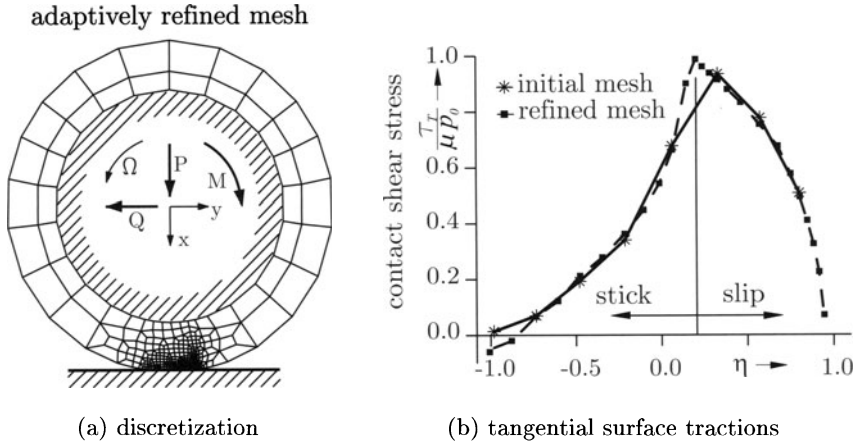
Fig. 1. Discretization and surface tractions for tractive rolling of an elastic ring [10]

simplified models and used a displacement oriented description of sticking and sliding contact. An early important contribution treating rolling contact by numerical methods has been published by Padovan [14]. Other publications in the vicinity of the authors are e.g. the PhD-thesis of Laging [6] and Gall [2]. From the present point of view the treatment of the contact conditions of the cited publications did not correctly take into account the strain gradient dependence of the tangential velocities of the contacting bodies. As far as the authors could find the PhD-thesis of Nackenhorst [8] was the first publication introducing the rolling contact condition to give exact conditions for sticking and sliding.

(a) adaptively refined finite element mesh

(b) pressure distribution of the initial and the refined discretization
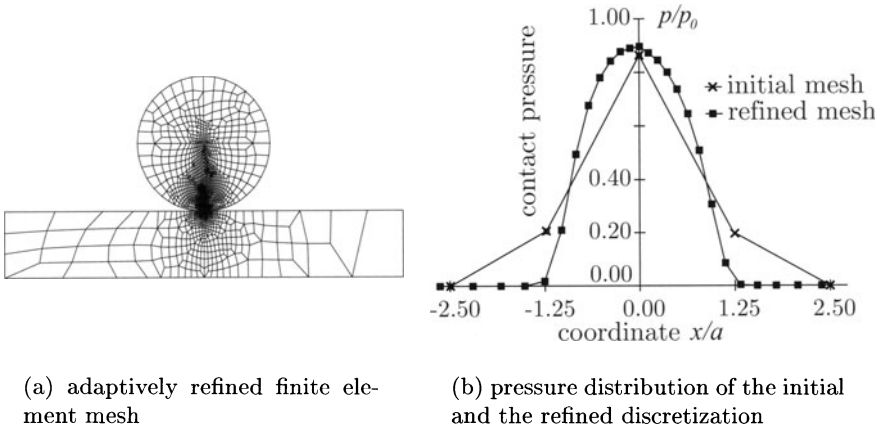
Fig. 2. Contact of an elastic disc with a rough elastic foundation [11]

As usual in those days the computation of contact was confined to so called elastic-rigid contact or contact of an elastic body with a rigid obstacle. For the treatment of rolling contact of a wheel on a rail several improvements of the contact description and the computer code had to be performed. Without going into detail these improvements have been:

- introducing an adaptive refinement for an efficient detection of the contact area including a distinction of sticking and sliding (see e.g. Fig. 1)
- extension from elastic-rigid to elastic-elastic contact
- extension from two dimensional to three dimensional simulation with an improved contact algorithm including the utilization of surface roughness (see e.g. Fig. 2 and Willner [16])

Both figures underline the importance of an adequate modelling of contact including the treatment of sticking and sliding. The following chapters provide the theoretical background for an effective evaluation of the rolling contact.

2   The Arbitrary LAGRANGian-EULERian Approach for Rolling Contact

Following the idea of a classical finite element approach for nonlinear contact problems the undeformed structure has to be discretized and depending on the required accuracy an adaptive refinement in the area of stress concentrations has to be performed. In the case of static conditions or dynamic applications without large rigid body motions one discretization for each load case is appropriate. If one would apply this approach for rolling contact a couple of difficulties would arise. First, the wheel would have to be discretized on its complete circumference with a fine mesh since for high velocity a subsequent refinement and coarsening of the contact area is numerically not efficient. Second, even for stationary rolling the time integration of the rolling process until a stable solution is reached requires a very long part of the rail to be discretized as well. Third, an extremely small time step has to be chosen for a reasonable integration taking into account the fine discretization of the contact zone. At the end it is uncertain whether the obtained solution is still reliable or whether the numerical errors might dominate.

An alternative method for the numerical simulation of problems with time dependent boundary conditions is the nonmaterial discretization using the Arbitrary LAGRANGian-EULERian approach (or abbreviated the ALE-approach) according to Nackenhorst [8] and [9]. The idea of this approach is to decompose the total deformation of the rolling wheel into a rigid body motion and a superimposed motion of the endpoint of the position vector into its final position. The first part is equivalent to an EULERian description since the material is moving through the discretized area as in computational fluid dynamics. The latter motion can be interpreted as a deformation of a material body using the LAGRANGian description.
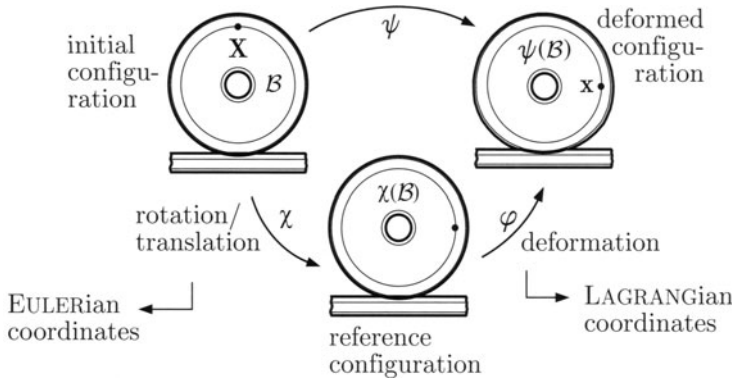
Fig. 3. ALE decomposition of the motion

Since the unknown deformation φ is relative to the EULERian reference configuration χ, this approach is sometimes called a relative kinematical description. In the stationary case χ is known a priori and φ is the only unknown field variable in the total deformation ψ. The decomposition can be written as the mapping of φ applied to χ:

$$\psi = \varphi \circ \chi. \tag{1}$$

Thus for the material time derivative also convective parts have to be considered:

$$\frac{d\,(\ldots)}{dt} = \left.\frac{\partial\,(\ldots)}{\partial t}\right|_{\chi} + \mathrm{Grad}\,(\ldots) \cdot \dot{\chi}. \tag{2}$$

The velocity of a material particle in the relative kinematical description reads as

$$\mathbf{v}\,(\mathbf{x},\,t) = \mathbf{v}\,(\varphi,\,t) = \left.\frac{\partial\varphi}{\partial t}\right|_{X} = \underbrace{\left.\frac{\partial\varphi}{\partial t}\right|_{\chi}}_{\text{relative velocity}} + \underbrace{\mathrm{Grad}\,\varphi \cdot \dot{\chi}}_{\text{convective velocity}}, \tag{3}$$

with the relative velocity describing the velocity of spatial points in relation to the reference configuration. The rigid body velocity of the wheel turning with an angular velocity $\boldsymbol{\Omega}$ is

$$\dot{\chi} = \left.\frac{\partial\chi}{\partial t}\right|_{X} = \left.\frac{\partial\chi}{\partial t}\right|_{\chi} + \boldsymbol{\Omega} \times \chi. \tag{4}$$

For further details concerning the ALE-kinematics of rolling contact it is referred to [13].
The so-called ALE-observer can be interpreted as travelling on the axle of the rolling wheel, not observing the "history of a material particle", but the motion of a particle at a certain place $\mathbf{x}$.

Advantages of this method are:

- stationary rolling is described independent of time
- a fine mesh zone only has to be concentrated to the contact regions which significantly decreases the number of degrees of freedom
- error estimation can be applied with respect to the spatial discretization only
- transient dynamics is described in space directly

Using this description difficulties arise in the case of "material with memory" and the formulation of frictional contact conditions.

3 Equations of Motion

For a sufficient mathematical model one has to start from the equation of motion

$$\text{Div}\,\hat{\mathbf{P}} + \hat{\varrho}\,\mathbf{b} = \hat{\varrho}\,\frac{d\mathbf{v}}{dt} \tag{5}$$

in the reference configuration with $\hat{\mathbf{P}}$ being the first PIOLA-KIRCHHOFF stress tensor and $\hat{\varrho}\,\mathbf{b}$ being the imposed loads. Through the evaluation of d'ALEMBERT's principle and the consideration of the dynamical boundary conditions one can obtain the weak formulation of the equation of motion for the rolling contact problem

$$\underbrace{\int\limits_{\chi(\mathcal{B})} \hat{\varrho}\frac{d\mathbf{v}}{dt}\cdot\boldsymbol{\eta}\,d\hat{v}}_{\text{virt. work of inertia forces}} + \underbrace{\int\limits_{\chi(\mathcal{B})} \hat{\mathbf{P}}\cdot\cdot\,\text{Grad}\,\boldsymbol{\eta}\,d\hat{v}}_{\text{virt. work of internal forces}} = \underbrace{\int\limits_{\chi(\mathcal{B})} \hat{\varrho}\mathbf{b}\cdot\boldsymbol{\eta}\,d\hat{v} + \int\limits_{\partial_t\chi(\mathcal{B})} \overline{\mathbf{T}}\cdot\boldsymbol{\eta}\,d\hat{a}}_{\text{virt. work of body and surface forces}},$$

$$\tag{6}$$

if the unknown contact stresses are treated as imposed surface loads. If the contact stresses are calculated as considered in Sec. 4, eq. (6) takes the form

$$\int\limits_{\chi(\mathcal{B})} \hat{\varrho}\frac{d\mathbf{v}}{dt}\cdot\boldsymbol{\eta}\,d\hat{v} + \int\limits_{\chi(\mathcal{B})} \hat{\mathbf{P}}\cdot\cdot\,\text{Grad}\,\boldsymbol{\eta}\,d\hat{v} - \int\limits_{\chi(\mathcal{B})} \hat{\varrho}\mathbf{b}\cdot\boldsymbol{\eta}\,d\hat{v} - \int\limits_{\partial_t\chi(\mathcal{B})\backslash\partial_c\psi(\mathcal{B})} \overline{\mathbf{T}}\cdot\boldsymbol{\eta}\,d\hat{a}$$

$$-\delta\int\limits_{\partial_c\psi(\mathcal{B})} (p\,d + \boldsymbol{\tau}\cdot\mathbf{s})\,da = 0. \tag{7}$$

The surface is now divided into the contact area $\partial_c\psi\,(\mathcal{B})$ with unknown contact stresses and the remaining surface $\partial_t\chi\,(\mathcal{B})\backslash\partial_c\psi\,(\mathcal{B})$ with prescribed stresses or displacements. The test function $\boldsymbol{\eta}$ can be interpreted as virtual displacement, which has to fulfill the geometrical boundary conditions. It has to be mentioned that within the ALE frame of reference $\boldsymbol{\eta}$ has the character of a spatial variation: the position $\boldsymbol{\varphi}$ (see Fig. 3) which the particles are

currently passing is varied. The dimension of the prospected contact area $\partial_c \psi(\mathcal{B}) \subset \partial \psi(\mathcal{B})$ in the deformed configuration in not known in advance, but can be determined by solving the contact problem according to Sec. 4. Equation (7) is valid for stationary as well as for transient rolling.

Within the ALE frame of reference special attention has to be paid to the calculation of the virtual work of the inertia forces, because the material derivative of the velocity according to eq. (2) is also considering convective parts:

$$\int_{\chi(\mathcal{B})} \hat{\varrho} \frac{d\mathbf{v}}{dt} \cdot \boldsymbol{\eta} \, d\hat{v} = \int_{\chi(\mathcal{B})} \hat{\varrho} \left(\left. \frac{\partial \mathbf{v}}{\partial t} \right|_\chi + \mathrm{Grad} \, \mathbf{v} \cdot \dot{\boldsymbol{\chi}} \right) \cdot \boldsymbol{\eta} \, d\hat{v}. \tag{8}$$

Evaluating this one obtains

$$\int_{\chi(\mathcal{B})} \hat{\varrho} \frac{d\mathbf{v}}{dt} \cdot \boldsymbol{\eta} \, d\hat{v} = -\int_{\chi(\mathcal{B})} \hat{\varrho} \left(\mathrm{Grad} \, \boldsymbol{\varphi} \cdot \dot{\boldsymbol{\chi}} \right) \cdot \left(\mathrm{Grad} \, \boldsymbol{\eta} \cdot \dot{\boldsymbol{\chi}} \right) \, d\hat{v} + \underbrace{\int_{\partial \chi(\mathcal{B})} \hat{\varrho} \boldsymbol{\eta} \cdot \left(\mathrm{Grad} \, \boldsymbol{\varphi} \cdot \dot{\boldsymbol{\chi}} \right) \dot{\boldsymbol{\chi}} \cdot \hat{\mathbf{n}} \, d\hat{a}}_{\text{momentum flux}}$$

$$\tag{9}$$

for the stationary case due to $\frac{\partial \boldsymbol{\varphi}}{\partial t} = 0$, with $\hat{\mathbf{n}}$ being the normal vector outward on $\partial \chi(\mathcal{B})$ in the reference configuration. For transient rolling it is referred to Sec. 6.1.

4   Description of Contact

For the evaluation of contacting bodies the weak formulation of the equation of motion (7) has to be considered for each of them and the boundary conditions have to be fulfilled, which is also the case for the contact zone. A priori neither the contact area nor the contact stresses are known, so first the detection of normal contact is required.

4.1   The Normal Contact Problem

The normal contact problem is characterized by the prohibition of the penetration d of the contacting bodies

$$d \leq 0 \tag{10}$$

and the fact that only compressive stresses

$$p \geq 0 \tag{11}$$

can occur in the contact zone. Because the inequalities (10) and (11) exclude each other, the SIGNORINI condition

$$p \, d = 0 \tag{12}$$

can be introduced. Within a weak formulation eq. (12) can be transformed
into

$$\delta \int_{\partial_c \psi(\mathcal{B})} p \, d \, da = 0 \tag{13}$$

as written in eq. (7). In the literature several methods are proposed to fulfill
this condition either exactly or approximately. Fig. 4 represents a graphical
representation of commonly used approaches.

constitutive relations following
a rough surface concept:
- ■ experimental data from Willner
- + experimental data from Knothe/Theiler

fictitious height of
the asperities
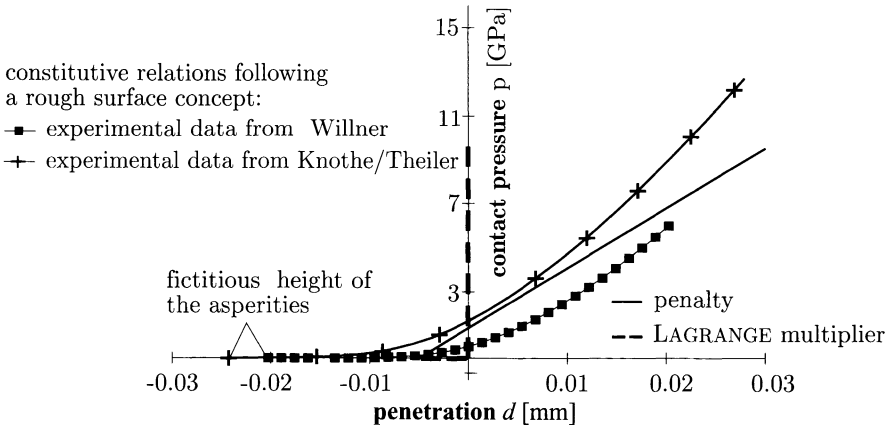
Fig. 4. Constitutive relations for normal contact [11]

The method of LAGRANGE multipliers would lead to an exact solution of the
prohibition of the penetration, but also causes some numerical difficulties.
Another idea is the introduction of the rough surface concept e.g. accord-
ing to Wriggers [18], thus introducing a meso-scale constitutive law which
requires a statistical description of the surface. The evaluation of the corre-
sponding statistical formulation of Willner [16] leads to nonlinear pressure-
penetration functions. Fig. 4 contains relations provided by Willner [16] and
Knothe/Theiler [4].
Another appropriate means is the introduction of the penalty method using
a linear relationship between the contact pressure p and the penetration d

$$p = \alpha_N d \tag{14}$$

with the stiffness parameter α_N. The weak formulation of the normal contact
condition (13) is now

$$\delta \int_{\partial_c \psi(\mathcal{B})} p \, d \, da \approx \delta \int_{\partial_c \psi(\mathcal{B})} \frac{\alpha_N}{2} d^2 \, da = \int_{\partial_c \psi(\mathcal{B})} \alpha_N d \, \delta d \, da. \tag{15}$$

For $\alpha_N \to \infty$ one would obtain an exact solution, but it is numerically unpractical to introduce the penalty factor going to infinity. Thus a certain penetration $d > 0$ of the contacting bodies has to be tolerated.
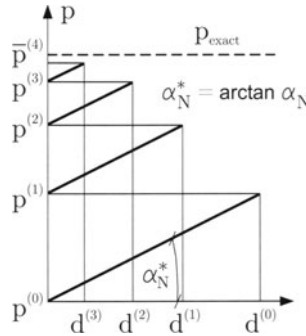
Fig. 5. Augmented Lagrangian formulation for the normal contact (Uzawa-algorithm) [13]

For a more accurate solution without an increase of the penalty factor the Augmented Lagrange approach can be introduced according to Fig. 5, see also [15]. The violation of the geometrical constraints is minimized in an iterative way by using

$$\delta \int_{\partial_c \psi(\mathcal{B})} p\, d\, da \approx \int_{\partial_c \psi(\mathcal{B})} \left(\overline{p}^{(k)} + \alpha_N\, d^{(k)} \right) \delta d\, da. \tag{16}$$

4.2   The Tangential Contact Problem for Rolling Contact

The tangential contact is described on the basis of the sticking condition and the friction law, which is used for the distinction of sticking and sliding:

$$\begin{array}{cccc}
 & \text{friction law} & \text{slip} & (17) \\
\text{sticking} & \mathcal{R} = \|\boldsymbol{\tau}\| - \tau_{max} \leq 0 & \|\mathbf{s}\| = \mathbf{0} & \\
\text{sliding} & \mathcal{R} = \|\boldsymbol{\tau}\| - \tau_{max} = 0 & \|\mathbf{s}\| \geq \mathbf{0} &
\end{array}$$

The maximum tangential shear stresses τ_{max} can be computed by an arbitrary constitutive law.

As mentioned in Sec. 2, within the ALE frame of reference the "history of a material particle" is not known a priori. As the relative displacement of two contacting particles therefore cannot be obtained directly, the slip vector $\mathbf{s}$ has to be replaced by the slip velocity $\dot{\mathbf{s}}$. Accordingly, the sticking condition for rolling contact is defined by

$$\|\dot{\mathbf{s}}\| = \mathbf{0}, \tag{18}$$

while in the case of sliding the slip velocity is given by

$$\|\dot{\mathbf{s}}\| = -\dot{\xi} \frac{\partial \mathcal{R}}{\partial \boldsymbol{\tau}}, \tag{19}$$

which describes the analogy between friction and non associated plasticity with $\dot{\xi}$ being a dimensionless parameter as well known from the literature. Similar to the normal contact problem the tangential contact conditions can now be defined by the KUHN-TUCKER condition

$$\|\dot{\mathbf{s}}\|\mathcal{R} = 0 \quad \wedge \quad \mathcal{R} \le 0 \quad \wedge \quad \|\dot{\mathbf{s}}\| \ge 0. \tag{20}$$

As shown in Fig. 6 within the ALE frame of reference the sticking condition
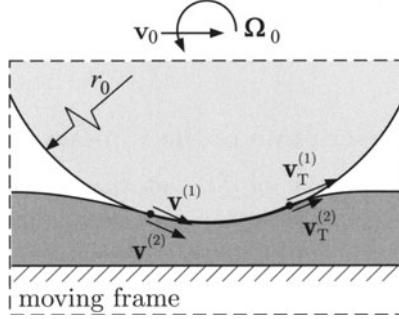
Fig. 6. Simplified sketch of contact requirements for the ALE formulation

requires identical velocities at the contact point while sliding requires the velocities of both contacting points lying in the common tangential plane of both surfaces.

Furthermore, the contact shear stresses are replaced by the corresponding force impact

$$\boldsymbol{\lambda} = \int_{t_1}^{t_2} \boldsymbol{\tau}\,dt \tag{21}$$

in order to obtain the weak formulation of the tangential contact problem

$$\underset{\partial_c\psi(\mathcal{B})}{\int} \boldsymbol{\tau}\cdot\delta\mathbf{s}\,da = \underset{\partial_c\psi(\mathcal{B})}{\int} \frac{d}{dt}\left(\boldsymbol{\lambda}\cdot\delta\mathbf{s}\right)da - \underset{\partial_c\psi(\mathcal{B})}{\int} \boldsymbol{\lambda}\cdot\delta\dot{\mathbf{s}}\,da, \tag{22}$$
$$\text{influx/outflux condition} \qquad \text{sticking condition}$$

which for stationary rolling in component notation turns into

$$\underset{\partial_c\psi(\mathcal{B})}{\int} \boldsymbol{\tau}\cdot\delta\mathbf{s}\,da = \underset{\partial_c\psi(\mathcal{B})}{\int} \tau_\alpha\delta\zeta^\alpha\,da = \underset{\partial_c\psi(\mathcal{B})}{\int} \frac{d}{dt}\left(\lambda_\alpha\delta\zeta^\alpha\right)da - \underset{\partial_c\psi(\mathcal{B})}{\int} \lambda_\alpha\delta\dot{\zeta}^\alpha\,da \tag{23}$$

with ζ^α as the convective coordinates of the slip vector $\mathbf{s}$, see Fig. 7. Similar to the normal contact problem this equation can be regularized using the described methods, but again it is advantageous to use the penalty method.

The slip velocity $\dot{\mathbf{s}}$ is additively divided into the "sticking velocity" $\dot{\mathbf{s}}_{st}$ and the "sliding velocity" $\dot{\mathbf{s}}_{sl}$, the sticking part is then treated as follows:

$$\dot{\mathbf{s}}_{st} = \dot{\mathbf{s}} - \dot{\mathbf{s}}_{sl} = -\frac{1}{\alpha_T}\boldsymbol{\lambda} \approx 0. \tag{24}$$

So in eq. (23) λ_α can be introduced as

$$\lambda_\alpha = -\alpha_T m_{\alpha\beta}\dot{\zeta}^\beta_{conv} \tag{25}$$

with $\dot{\zeta}^\beta_{conv}$ being the convective part of the "sliding velocity" and $m_{\alpha\beta}$ being the metric of the convective vectors (see Sec. 4.3). For transient contact problems see Sec. 6.

4.3   Kinematical Description of the Contact

For the kinematical description of the contact some special formulations are necessary. The calculational datum planes of roughness are considered as surfaces, which implies introducing a limit value of gap for the detection of contact. The gap or the penetration of two contacting particles is calculated
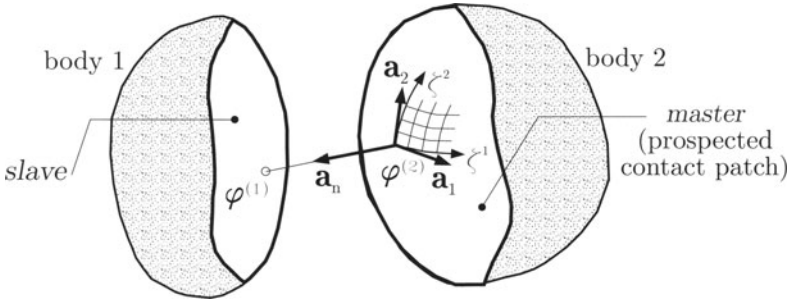
Fig. 7. Contact kinematics for elastic bodies (exploded view)

by introducing a parameterization $\boldsymbol{\Psi}$ of body 2 (*master body*) in the deformed configuration within the ALE representation as shown in Fig. 7:

$$\boldsymbol{\varphi}^{(2)} = \boldsymbol{\Psi}\left(\zeta^1, \zeta^2, t\right) = \boldsymbol{\Psi}_t\left(\zeta^1, \zeta^2\right). \tag{26}$$

The tangential and the normal vectors of this parameterization are defined by $\mathbf{a}_\alpha$ and $\mathbf{a}_n$, $m_{\alpha\beta}$ is the corresponding metric. A point on body 1 (*slave body*)

$$\boldsymbol{\varphi}^{(1)} = \hat{\boldsymbol{\psi}}^{(1)}\left(\boldsymbol{\chi}^{(1)}, t\right) \tag{27}$$

is observed and the corresponding target point $\overline{\boldsymbol{\varphi}}^{(2)} = \boldsymbol{\Psi}_t\left(\overline{\zeta}^1, \overline{\zeta}^2\right)$ is identified; the overline is introduced to characterize the target point. Minimizing the penetration function, the penetration

$$d = -\left(\boldsymbol{\varphi}^{(1)} - \overline{\boldsymbol{\varphi}}^{(2)}\right) \cdot \overline{\mathbf{a}}_n \tag{28}$$

and the slip velocities

$$m_{\alpha\beta}\dot{\zeta}^{\beta} = (\mathbf{v}_1 - \mathbf{v}_2) \cdot \bar{\mathbf{a}}_\alpha \tag{29}$$

can be calculated. For the variation of the penetration and the slip velocity, which are necessary for eq. (7), and for more details see [13].

5   The Finite Element Formulation (FEM)

Based upon the fundamentals of the previous sections, the transcription of the nonlinear problem into the FEM is carried out by applying a consistent linearization according to [17]. The motion is incrementally formulated

$$^{t+\Delta t}\varphi =\,^t \varphi + \Delta\varphi \tag{30}$$

by using an isoparametric approximation

$$\varphi\left(\chi,\, t\right) = \mathbf{H}\left(\chi\right)\hat{\varphi}\left(t\right) \tag{31}$$

for the spatial displacement field with $\hat{\varphi}\left(t\right)$ being the nodal values of the displacement field. Then the incremental finite element equation of motion for stationary rolling can be obtained:

$$\begin{bmatrix} {}^t\mathbf{K}^{(1)} - \mathbf{W}^{(1)} + {}^t\mathbf{C}_n^{(11)} + {}^t\mathbf{C}_t^{(11)} & {}^t\mathbf{C}_n^{(12)} + {}^t\mathbf{C}_t^{(12)} \\ {}^t\mathbf{C}_n^{(21)} + {}^t\mathbf{C}_t^{(21)} & {}^t\mathbf{K}^{(2)} - \mathbf{W}^{(2)} + {}^t\mathbf{C}_n^{(22)} + {}^t\mathbf{C}_t^{(22)} \end{bmatrix} \begin{bmatrix} {}^{t+\Delta t}\Delta\hat{\varphi}^{(1)} \\ {}^{t+\Delta t}\Delta\hat{\varphi}^{(2)} \end{bmatrix}$$

$$= \begin{bmatrix} {}^{t+\Delta t}\hat{\mathbf{f}}_e^{(1)} - {}^t\hat{\mathbf{f}}_\sigma^{(1)} + {}^t\hat{\mathbf{f}}_i^{(1)} - {}^t\hat{\mathbf{f}}_n - {}^t\hat{\mathbf{f}}_t \\ {}^{t+\Delta t}\hat{\mathbf{f}}_e^{(2)} - {}^t\hat{\mathbf{f}}_\sigma^{(2)} + {}^t\hat{\mathbf{f}}_i^{(2)} + {}^t\hat{\mathbf{f}}_n + {}^t\hat{\mathbf{f}}_t \end{bmatrix}, \tag{32}$$

here written for two contacting bodies $i = 1,\ 2$. The tangential stiffness matrix ${}^t\mathbf{K}^{(i)}$ is calculated according to Bathe [1]. Furthermore, the ALE inertia matrix

$$\mathbf{W}^{(i)} = \int_{\chi^{(i)}(\mathcal{B})} \hat{\varrho}^{(i)} \mathbf{A}^{(i)T} \mathbf{A}^{(i)} d\hat{v} \tag{33}$$

as well as the contact stiffness matrices can be obtained according to Nackenhorst [13]. In eq. (33) matrix $\mathbf{A}^{(i)}$ defines the relation of the convective velocity $\mathrm{Grad}\,\varphi^{(i)} \cdot \dot{\chi}^{(i)}$ and the nodal values of the displacement field $\hat{\varphi}^{(i)}$ by

$$\mathrm{Grad}\,\varphi^{(i)} \cdot \dot{\chi}^{(i)} = \mathbf{A}^{(i)}\hat{\varphi}^{(i)}. \tag{34}$$

The equivalent nodal forces of body and surface loads are contained in ${}^{t+\Delta t}\hat{\mathbf{f}}_e^{(i)}$ and ${}^t\hat{\mathbf{f}}_\sigma^{(i)}$ are the internal forces according to Bathe [1]. The equivalent inertia

forces, which are very important within the used ALE frame of reference, are defined as

$$
{}^{t}\hat{\mathbf{f}}_{i}^{(i)} = \int\limits_{\chi^{(i)}(\mathcal{B})} \varrho^{(i)} \mathbf{A}^{(i)T} \operatorname{Grad} \boldsymbol{\varphi}^{(i)} \cdot \dot{\boldsymbol{\chi}}^{(i)} d\hat{v}
\tag{35}
$$

and the equivalent contact forces as

$$
{}^{t}\hat{\mathbf{f}}_{n} = \int\limits_{\partial_{c}\psi(\mathcal{B})} \mathbf{H}^{T}\mathbf{a}_{n}\,{}^{t}p\,da \quad \text{and} \quad {}^{t}\hat{\mathbf{f}}_{t} = \int\limits_{\partial_{c}\psi(\mathcal{B})} \mathbf{H}^{T}\mathbf{a}_{\alpha}\,{}^{t}\tau_{\alpha}\,da.
\tag{36}
$$

The system of equations of motion (32) is then treated with an iterative algorithm, e.g. the NEWTON-RAPHSON scheme, to find the equilibrium solution for the system. The calculation of the contact itself is carried out within a loop over all elements of the slave group (see Sec. 4.3): the corresponding points of the master elements are identified and their contributions are calculated. For the distinction of sticking and sliding special attention has to be paid to the used ALE frame of reference. As introduced in eq. (21), the tangential contact stresses τ_{α} are substituted by the corresponding force impact λ_{α}. In the case of sticking the contact shear stresses are not calculated a priori. In [13] Nackenhorst therefore proposes an explicit predictor-corrector algorithm, where the force impact λ_{α}^{pred} is calculated in order to obtain the equivalent contact forces and contact stresses for the distinction of sticking and sliding.

6   Numerical Analysis of Transient Rolling Contact

6.1   Fundamentals for Transient Rolling

For the analysis of transient rolling contact problems it is important to consider the effects of the dynamical behavior of rotating bodies. Thus a full dynamical analysis of the problem is necessary. For transient rolling one can again start from the equation of motion (5) and obtain eq. (7), which is valid for all cases of non-stationary rolling. Special attention has to be paid to the contribution of the virtual work of the inertia forces (eq. (8)), where again the time derivative of the velocity (see eq. (3)) of a material particle in the relative kinematical basis is as follows:

$$
\frac{d\mathbf{v}}{dt} = \mathbf{a}\left(\boldsymbol{\varphi},\,t\right) = \left.\frac{\partial^{2}\boldsymbol{\varphi}}{\partial t^{2}}\right|_{\boldsymbol{\chi}} + 2\,\operatorname{Grad}\,\left.\frac{\partial\boldsymbol{\varphi}}{\partial t}\right|_{\boldsymbol{\chi}} \cdot \dot{\boldsymbol{\chi}} + \operatorname{Grad}\left(\operatorname{Grad}\,\boldsymbol{\varphi} \cdot \dot{\boldsymbol{\chi}}\right) \cdot \dot{\boldsymbol{\chi}}.
\tag{37}
$$

This equation contains gradients of second order for stationary as well as for non-stationary rolling (see eq. (8)). This demands a C^{1}-finite element approximation, but opposite to stationary rolling all contributions of eq. (8) have to be considered since $\frac{\partial\boldsymbol{\varphi}}{\partial t}$ is no longer equal to zero. Applying a transcription of this equation according to Nackenhorst [13] leads to a formulation

with gradients of first order only, which is sufficient for a C^0-finite element approximation,

$$\int_{\chi(\mathcal{B})} \hat{\varrho} \frac{d\mathbf{v}}{dt} \cdot \boldsymbol{\eta} \, d\hat{v}, = \int_{\chi(\mathcal{B})} \hat{\varrho} \left(\left. \frac{\partial \mathbf{v}}{\partial t} \right|_{\chi} \cdot \boldsymbol{\eta} - \mathbf{v} \cdot \operatorname{Grad} \boldsymbol{\eta} \cdot \dot{\boldsymbol{\chi}} \right) d\hat{v} + \underbrace{\int_{\partial \chi(\mathcal{B})} \hat{\varrho} \boldsymbol{\eta} \cdot \mathbf{v} \dot{\boldsymbol{\chi}} \cdot \hat{\mathbf{n}} \, d\hat{a}}_{\text{momentum flux}}.$$

(38)

For the contact term transformations are also necessary, which altogether can be transformed into an incremental finite element formulation for the transient rolling contact problem for each body i,

$${}^t\mathbf{M}^{(i)} \, {}^{t+\Delta t}\ddot{\hat{\boldsymbol{\varphi}}}^{(i)} + {}^t\mathbf{G}^{(i)} \, {}^{t+\Delta t}\dot{\hat{\boldsymbol{\varphi}}}^{(i)} +$$

$$+ \left[{}^t\mathbf{K}^{(i)} - \mathbf{W}^{(i)} + {}^t\mathbf{C}_n^{(ii)} + {}^t\mathbf{C}_t^{(ii)} \right] {}^{t+\Delta t}\Delta\hat{\boldsymbol{\varphi}}^{(i)} = {}^{t+\Delta t}\mathbf{R}^{(i)}$$

(39)

For the calculation of the matrices ${}^t\mathbf{K}^{(i)}$, $\mathbf{W}^{(i)}$ as well as for the contact stiffness matrices see Sec. 5, ${}^t\mathbf{M}^{(i)}$ is the mass matrix and ${}^t\mathbf{G}^{(i)}$ the gyroscopic matrix containing contribution from Coriolis forces. ${}^{t+\Delta t}\mathbf{R}^{(i)}$ contains all forces including the contact forces.

6.2   Classification of Irregularities and Mechanisms of Excitation

Transient rolling demands the development of new solution methods and in order to choose the appropriate calculation method the irregularities have to be classified. For the treatment within the FEM three kinds of irregularities are distinguished by characterizing the surface irregularities by the parameter $\overline{L} = \frac{L}{a}$ according to Fig. 8, which is the wavelength of the disturbance divided by a typical measure of the contact patch, e.g. the width a.
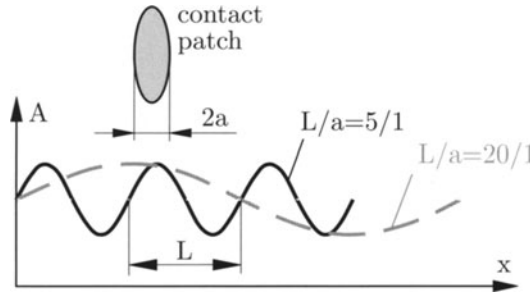
Fig. 8. Dimensions of irregularities and contact patch [5]

First, there are excitations in the high frequency range with $\overline{L} \leq 0.1$. This might be the case e.g. for travelling over surface roughness. The changes take place very fast and thus no reasonable macroscopic dynamic reactions can be identified. In this case the modal superposition is a promising approach, where the solution of the stationary rolling contact problem is frozen and taken as

initial configuration. Therefore only small relative deformations have to be considered.

The same is the case for the treatment of excitations in the low frequency range ($\overline{L} \geq 20$), which might appear when the wheel is travelling across sleepers, then the rail changes its curvature due to loading. It should be mentioned in this context, that for the application of the modal analysis it is necessary to use suitable eigenvalue solvers for the treatment of unsymmetric matrices which are currently developed.

But most of the difficulties arise when the size of irregularities is of the same order as the size of the contact patch ($0.1 \leq \overline{L} \leq 20$), e.g. in the case of short pitch corrugation. Massive dynamic reactions have to be noticed as well as varying wheel loads and therefore a time-dependent gap function, e.g.

$$g\left(t\right) = g_0 \left(1 - \cos \frac{2\pi}{L} V t\right) \tag{40}$$

can be introduced in order to simulate these irregularities. Typical values for rails with reasonable short pitch corrugation are $L = 2 - 10\,\mathrm{cm}$ for the wavelength, and $g_0 = 50 - 300\,\mu\mathrm{m}$ for the amplitude of the irregularity. Additional nonlinear reactions within the contact zone appear, which are of the same order of magnitude as the static approach of the contacting bodies. Thus direct time-step integration procedures are necessary. The well known NEWMARK scheme is proposed.

6.3   Transient Formulation of the Contact Conditions

For a transient analysis again the penalty approach according to Sec. 4.2 is used, the equation for the components of the force impact vector from eq. (25) now turns into

$$\lambda_\alpha = -\alpha_T m_{\alpha\beta} \left(\dot{\zeta}^\beta_{rel} + \dot{\zeta}^\beta_{conv}\right), \tag{41}$$

with $\dot{\zeta}^\beta_{rel}$ and $\dot{\zeta}^\beta_{conv}$ being the relative motion and the convective part of the slip velocity in the convective frame of reference respectively [13]. The values of ζ^β are calculated from eq. (29). The virtual work of the time dependent change of the tangential force impact in eq. (23)

$$\int_{\partial_c \psi(\mathcal{B})} \left.\frac{\partial \lambda_\alpha}{\partial t}\right|_\chi \delta\zeta^\alpha \, da = \int_{\partial_c \psi(\mathcal{B})} -\alpha_T m_{\alpha\beta} \left.\frac{\partial \dot{\zeta}^\beta}{\partial t}\right|_\chi \delta\zeta^\alpha \, da \tag{42}$$

has to be considered. The contact stiffness matrices and the equivalent contact forces are now calculated as discussed in Sec. 5. It should be mentioned that the contacting bodies are now coupled by the mass and gyroscopic matrices, unsymmetric matrices are the consequences which have to be treated with implicit integration procedures.

7 Numerical Results

Numerous parameter studies have been carried out during the project for free rolling in order to obtain the contact stresses in normal and tangential direction. These data are fundamentals for the investigation of friction and wear. Here the results for a wheel S1002 and a rail UIC60 shall be provided.

view of half of the discretization

12052 nodes
54389 tetrahedron elements
3072 contact elements

wheel: S1002

rail: UIC60
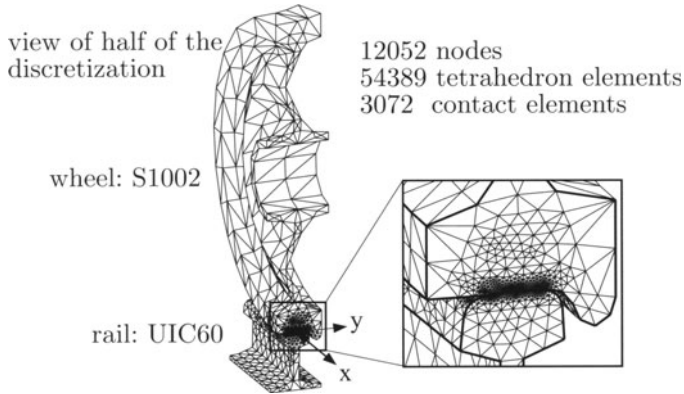
Fig. 9. FE discretization of the wheel-rail contact [12]

The wheel has a diameter of approximately $1.2\,m$, the contact zone has a size of approximately $2\,cm^2$. A suitable finite element discretization as shown in Fig. 9 is chosen. Because of the used ALE frame of reference the fine mesh can be concentrated to the explicit prospected contact region (see Sec. 2). Details of the discretization are shown in Fig. 9.

Linear elastic material behavior has been assumed with $E = 210\,GPa$ and $\nu = 0.3$. The contact conditions of wheel and rail are characterized by the geometry at the contact point. In general, the lateral wheel position on the rail is not fixed to the center line. Thus parameter studies for different laterally
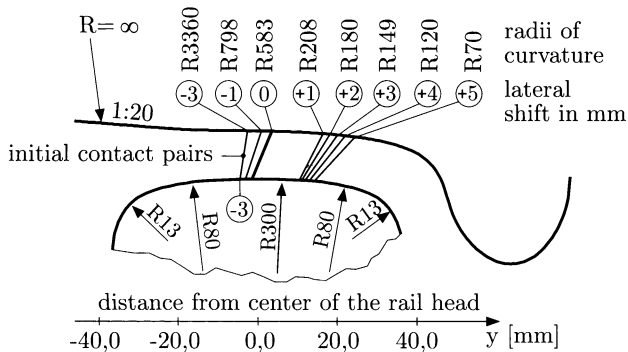
Fig. 10. Parameter definition of the contact setup, radii of curvature
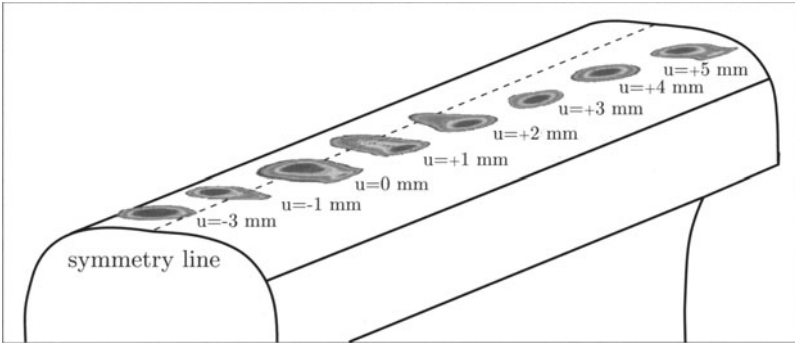
Fig. 11. Comparison of the position of the contact patches dependent on the variation of the lateral shift

lateral shift	$p^{(90)}$	$\sigma_V^{(90)}$	$p^{\text{Hertz}(90)}$
[mm]	[MPa]	[MPa]	[MPa]
-3,0	824,6	442,9	1243,8
-1,0	849,6	490,1	989,3
0,0	809,7	441,2	981,2
1,0	549,3	302,2	———
2,0	675,2	385,3	1520,0
3,0	840,9	481,3	1575,8
4,0	865,8	488,6	1448,4
5,0	630,3	323,6	———

Table 1. Scaled values of stresses for free rolling

shifted positions of the wheel as shown in Fig. 10 were carried out. For a velocity $v = 200\,km/h$ and a wheel load of approx. $90\,kN$ the contact patches are shown in the overview in Fig. 11. The maximum values for the contact pressure and the v. MISES stresses are compiled in table 1, which additionally contains the corresponding HERTZian solutions.
The following can be noticed:

- due to the application of a nonlinear finite element model the numerical solutions differ significantly from the HERTZian solution
- the maximum contact pressure is overestimated by the HERTZian theory
- for some constellations no corresponding HERTZian solution can be found due to the curvature at the initial contact points

This can be explained by the nonlinear FEM which takes into account all geometrically nonlinear effects, e.g. the local change of the radii of curvature at the contacting points due to loading. The convergence of these radii e.g. leads to larger contact patches and smaller maximum contact pressures. The same is the case due to decreased stiffness in the contact area caused by the considered surface roughness.
It has to be noticed that between a lateral shift of $-1\,mm$ and $+2\,mm$ a change of the point of the maximum contact pressure takes place. The "main
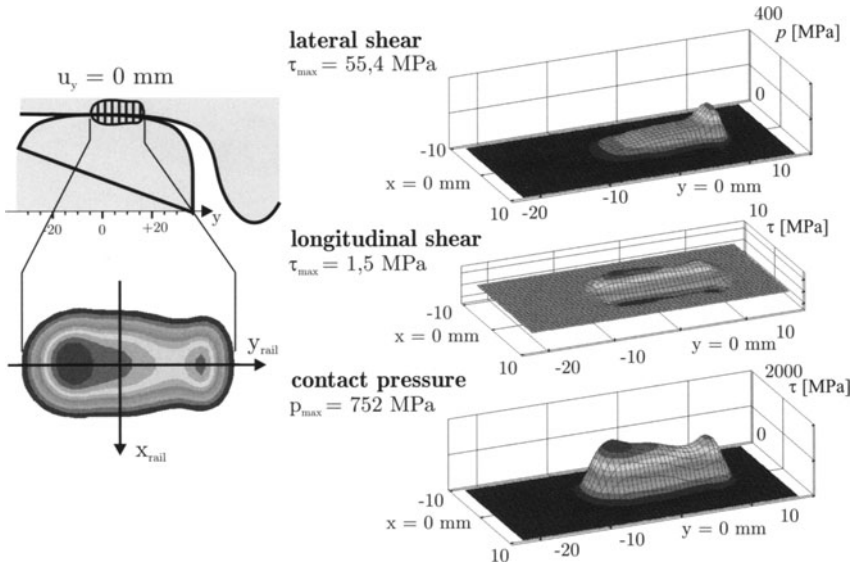
Fig. 12. Surface traction distribution for an unshifted situation (lateral shift $\pm 0\,mm$)

contact point" moves to an area with a different curvature, so for a situation without a lateral shift two maxima of contact pressure occur as shown in Fig. 12. For a lateral shift of $+5\,mm$ the same seems to be the case, and it has to be mentioned, that due to the fact of the curvature of the wheel being larger than the curvature of the rail no corresponding HERTZian contact ellipse can be found.

Furthermore, in Fig. 12 and 14 the distributions of lateral and longitudinal shear stresses are given, where the longitudinal stresses show the expected antimetric shape, but the values seem to be negligibly small.
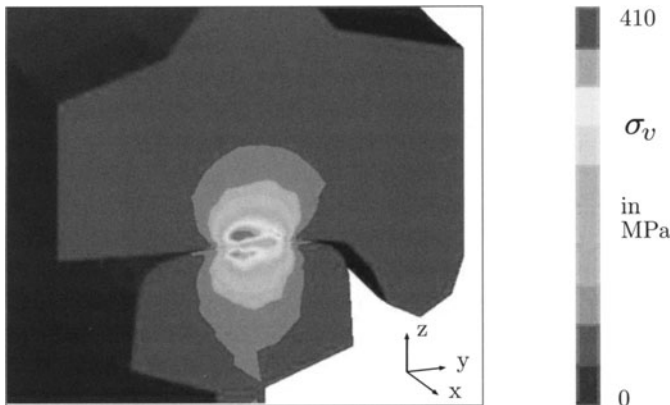
Fig. 13. v. Mises stresses (σ_V) for a lateral shift $\pm 0\,mm$
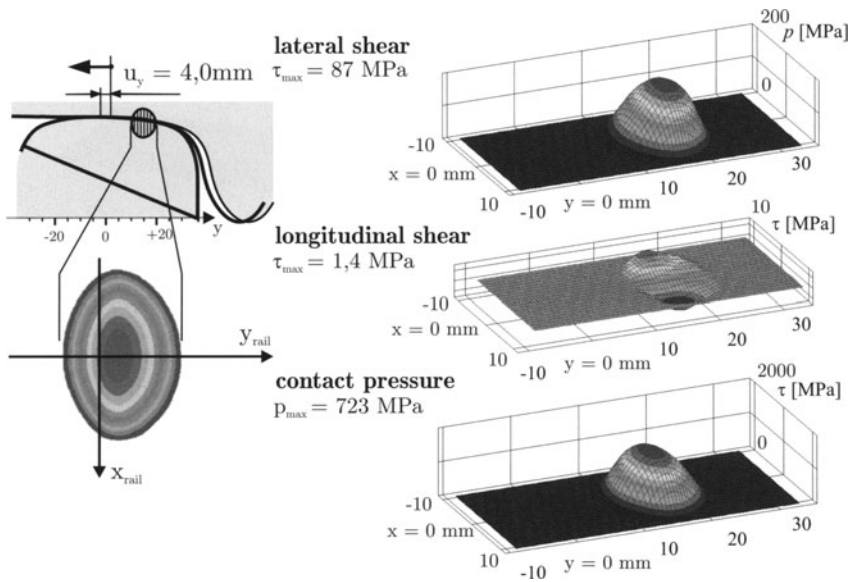
Fig. 14. Surface traction distribution for a lateral shift $+4\,mm$

With a coefficient of friction of $\mu = 0,15$, full sticking can be assumed for all analyzed cases. The V. MISES stresses for the lateral shift of $\pm 0\,mm$ are shown in Fig. 13. As expected for the traction free case the maximum appears inside the bodies at a location near to the surface.
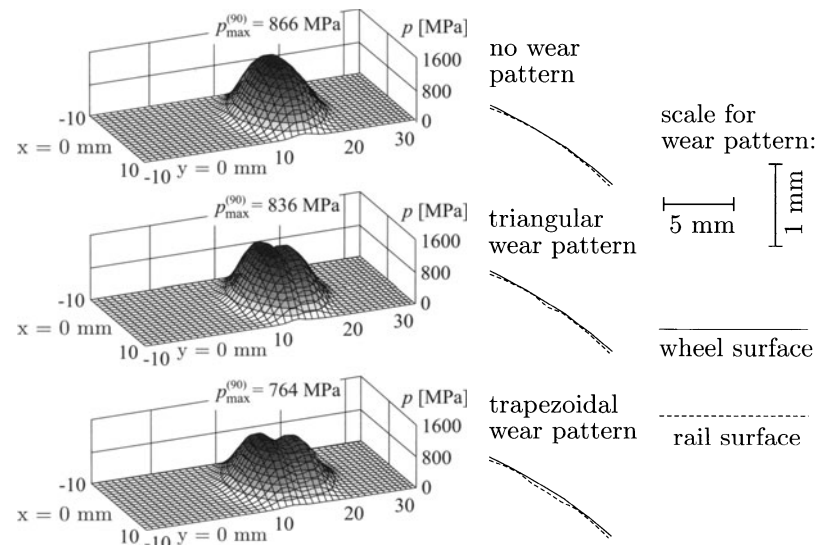
Fig. 15. Comparison of the contact pressure distribution for different worn profiles

For a lateral shift of $+3\,mm$ and $+4\,mm$ one obtains solutions which are quite similar to the HERTZian solution, which corresponds well with the results provided by Le The [7]. The contact stress distribution for $+4\,mm$ is given in Fig. 14. The results for this lateral shift also cover the influence of generic wear patterns of triangular shape and trapezoidal shape, see Fig. 15. Since the position of the maximum wear has been assumed under the position of the maximum pressure, the contact area will increase thus reducing the maximum contact pressure. But this is only a numerical result for this special setup and can not be generalized.

To model the real contact it will be necessary to take into account additional surface irregularities such as short pitch corrugation, as mentioned in Sec. 6.

8   Conclusion and Outlook

The presented numerical results for the rolling contact situation underline the importance of an adequate modelling of the correct contact situation since a shift of the center plane of only $1\,mm$ can change the contact patch as well as the distribution of stress and strain in the vicinity of the contact patch significantly. So far only free rolling on a straight track has been analyzed thoroughly. Since no constitutive relation for the dependence of the shear stresses from the slip velocity is available, up to now only sticking using the penalty approach has been simulated. The program code nevertheless is capable for the superposition of spin motion and two point contact.

In cooperation with the other members of the working group "contact, friction and wear" a constitutive model for tangential contact conditions of rough surfaces and a material description taking into account the plastic behavior of the used steel shall be developed next. With sufficient progress in these fields, additional parameter studies of more advanced running situations will be performed.

The transient rolling situation requires intensive additional research as well as considerations about a reliable error analysis and an efficient adaptive finite element technique. In the next years the provided method can economically be applied for specific contact situations only. Nevertheless this is an appropriate means to explore the confidence range of various approximate methods for the real time analysis of running trains.

Acknowledgement

References

1. Bathe, K.-J. (1990) Finite-Elemente-Methoden. Springer-Verlag, Berlin

2. Gall, R. (1985) Zur Berechnung von faserverstärkten Reifen mit der Methode der Finiten Elemente. Dissertation, Institut für Mechanik, Universität der Bundeswehr Hamburg

3. Kalker, J. J. (1990) Three–Dimensional Elastic Bodies in Rolling Contact. Kluver, Dordrecht

4. Knothe, K., Theiler, A. (1997) Rauhigkeitsmeßdaten betriebsbeanspruchter Schienen. Interner Bericht der Arbeitsgruppe Gleis. Technische Universität Berlin

5. Knothe, K., Wille, R., Zastrau, B. (2001) Advanced Contact Mechanics - Road and Rail. International Journal of Vehicle Mechanics and Mobility. **35**, No. 4-5, 361–407

6. Laging, G. (1985) Zur quasistatischen Berechnung des Deformationsverhaltens von Reifen. Dissertation, Institut für Mechanik der Hochschule der Bundeswehr Hamburg

7. Le The, H. (1987) Normal- und Tangentialspannungsverteilung beim rollenden Kontakt für Rotationskörper mit nichtelliptischen Kontaktflächen. Dissertation, TU Berlin

8. Nackenhorst, U. (1992) Zur Berechnung schnell rollender Reifen mit der Finite Element Methode. Dissertation, Institut für Mechanik, Universität der Bundeswehr Hamburg

9. Nackenhorst, U. (1993) On the finite element analysis of steady state rolling contact. In: ALIABADI, M. H.; BREBBIA, C. A.: *Contact Mechanics I – Computational Techniques*, Computational Mechanics Publication. 53–60

10. Nackenhorst, U. (1995) An adaptive finite element method to analyse contact problems. In: ALIABADI, M. H.; ALESSANDRI, C.: *Contact Mechanics II – Computational Techniques*, Computational Mechanics Publication. 241–248

11. Nackenhorst, U., Zastrau, B. (1997) Rolling contact of elastic bodies - Basic aspects of a finite element approach. In: MESKOURIS, K.; WITTEK, U.: *Aspects in Modern Computational Structural Analysis*. 385–400

12. Nackenhorst, U., Zastrau, B., Jarewski, J. (2000) Finite Element Modeling of 3D Elastic-Elastic Rolling Contact. ZAMM. **8**, suppl. 1, 57–60

13. Nackenhorst, U. (2000) Rollkontaktdynamik: Numerische Analyse der Dynamik rollender Körper mit der Finite Element Methode. Habilitation, Institut für Mechanik, Universität der Bundeswehr Hamburg

14. Padovan, J. (1987) Finite element analysis of steady-state and transiently moving/rolling nonlinear viscoelastic structures. Computers & Structures. **27(2)**, 249–257

15. Simo, J. C., Laursen, T. A.. (1992) An augmented Lagrangian treatment of contact problems involving friction. Computers & Structures. **42 (1)**, 97–116

16. Willner, K. (1995) Ein statistisches Modell für den Kontakt metallischer Körper. Dissertation, Institut für Mechanik, Universität der Bundeswehr Hamburg

17. Wriggers, P. (1988) Konsistente Linearisierungen in der Kontinuumsmechanik und ihre Anwendung auf die Finite Element Methode. Habilitation, Universität Hannover

18. Wriggers, P. (1995) Finite Element Algorithms for Contact Problems. Arch. Comp. Meth. Engg., **2**, No. 4, 1–49

Experimental Analysis of the Cyclic Deformation and Damage Behavior of Characteristic Wheel and Rail Steels

Bernhard Denne[1], Karl-Heinz Lang[2], and Detlef Löhe[2]

[1] Institut für Werkstoffkunde I, Universität Karlsruhe (TH), now with Bosch GmbH, 77813 Bühl, Germany
[2] Institut für Werkstoffkunde I, Universität Karlsruhe (TH), Kaiserstrasse 12, 76131 Karlsruhe, Germany

Abstract. As part of an extensive study exemplary results of the basic examinations which were carried out within this project, transmission electron microscopic (TEM) investigations of the microstructure of a cyclic loaded rail steel and the deformation and lifetime behavior of a cyclic loaded wheel steel are represented. For the TEM investigations specimens are worked out of the head of unused rails. With these specimens push-pull tests under several loading conditions are carried out. After failure foils are taken from the gauge length to perform TEM investigations to determine the appearing dislocation structure in the ferritic phase of the pearlite. Additionally the observed dislocation structures are compared with that one found nearby the surface of a used rail taken from a high speed track and that one found in unused rails in the as-received condition. To investigate the cyclic deformation and lifetime behavior of a wheel steel specimens were worked out from several layers of the rim of a monobloc wheel, grade R7, used in high speed trains. Due to the heat treatment at the end of the manufacturing process there are different microstructures within several regions of the wheel rim. The specimens represent these different microstructures. Stress controlled fatigue tests are performed with these specimens. The cyclic deformation behavior, the lifetime behavior and the endurance limit are deduced from push-pull tests. The influence of the different microstructures on the deformation and lifetime behavior is discussed.

1 Scope of the Project

The intention of the project is to improve the basic understanding of the material behavior under the specific loading conditions of the system wheel/rail on the one hand and to enhance substantially the description of the loading capacity of the concerning materials on the other hand. With that, the still existing gaps in the knowledge on the failure mechanisms of rolling contact fatigue should be reduced and the conditions for a safe and economical dimensioning of components of the wheel/rail system should be improved. By comparison of the microstructure arising in defined loaded specimens with the one found in the highest loaded areas of various service loaded components conclusions on the loadings which actually appeared in operation get possible. With that, the possibility arises to correlate the results from fundamental

or mechanisms orientated experiments with the results from technological or lifetime oriented experiments as well as with the knowlcdge from the examination of service loaded components. The research project concentrates on two primary objectives:

- By specific light and electron microscopic examinations of the high loaded areas of used wheels and rails the microstructure resulting from the service load was determined. Typical areas of wheels and rails which were rejected from service because of macroscopic damages were analyzed. Parallel to this, specimens which were taken from new wheels and rails of the corresponding materials were investigated after various cyclic loadings. Thereby, the imposed loading was varied in steps from uniaxial homogeneous via uniaxial inhomogeneous to multiaxial inhomogeneous. Supplementary, rolls which were loaded in a special rolling testing device at the Institut für Maschinenkonstruktionslehre at the Universität Magdeburg were included into the examinations. The microstructural states of these specimens were compared with the ones found in the service loaded components.
- The cyclic deformation and failure behavior of typical materials states Of characteristic wheel and rail materials was determined using different loading conditions. At first, the material reaction and the resulting microstructure was determined using uniaxial homogeneous tension-compression loading without and with superimposed mean loads. Thereby, nominal stress control as well as total strain control was realized. After that as intermediate steps experiments with inhomogeneous loading (cyclic bending) and biaxial loading (cyclic torsion) were carried out. Finally, with an inhomogeneous multiaxial loading (local Hertzian pressure without and with additional shear stresses in the contact zone) the service loading conditions appearing in the rolling wheel/rail contact was further approximated. The damage and the lifetime behavior as well as the developing microstructure were identified for all loading conditions investigated.

In the following transmission electron microscopic investigations of the microstructure of a cyclic loaded rail steel (Section 2) and the deformation and lifetime behavior of a cyclic loaded wheel steel (Section 3) are represented as some exemplary results of the basic examinations. More information especially concerning the cyclic deformation, the damage and the lifetime behavior of a rail steel are included in the publications [1-7]. The experiments with Hertzian pressure and additional sheer loading in the contact zone are not completed yet. The results of these experiments will be published at a later time together with some finite elements calculations which still are carried out at the Institut für Baumechanik und Bauinformatik of the Universität of Dresden.

2   TEM Investigations of the Microstructure of a Cyclic Loaded Rail Steel

2.1   Materials and Testing Specimens

The 60 kg/m rails specified as "UIC 60 900 A" and used in these investigations are made of a standard carbon steel and were produced by Thyssen Stahl AG in conformity with UIC 860 V. One of them has been used in a commercial high speed railway line having a mixed accumulated service tonnage consisting of high speed and heavy load traffic. The other one was unused in the as received condition. Samples of the used rails were collected from the regions being highly stressed by rolling contact. From the unused rails samples have been taken from similar regions. Additionally solid round specimens were worked out from the head of unused rails. With these specimens total strain- controlled fatigue tests were performed at room temperature using a Schenck 100 kN servohydraulic testing machine. These tests were performed at $R_e = -1$ with a triangular waveform, at total strain amplitudes of 0.18 % and 1 %. Under total strain control the fatigue limit at a survival probability of 50 % is 0.18 % [4], determined by the staircase method [8]. The specimen investigated has reached the ultimate number of cycles ($N = 2 \cdot 10^6$) without failure. At $\varepsilon_{a,t} = 1.0\%$ the specimen failed at $N_f = 831$ cycles. These two amplitudes were chosen to generate different amounts of plastic deformation. In general the deformation behavior of the rail steel is characterized by cyclic softening followed by cyclic hardening as it is typical for normalized ferritic-pearlitic microstructures [6,9]. At $\varepsilon_{a,t} = 1.0$ % the maximum plastic strain amplitude was 0.66 %. Remarkable is the fact that even at $\varepsilon_{a,t} = 0.18$ % where the specimen did not fail within $2 \cdot 10^6$ cycles a maximum of the plastic strain amplitude of $\varepsilon_{a,p} = 0.01$ % was found [9].

For TEM analyses, the gauge sections of the fatigued specimens were sliced along the transversal direction. All the TEM specimens were prepared by mechanical thinning and then jet polishing using a Struers, "A8", electrolite. The chemical composition and mechanical properties of the rail steel are summarized in Tabels 1 and 2. The steel has a fully pearlitic microstructure with MnS inclusions elongated along the rolling direction due to the hot rolling during the manufacturing process. The pearlite structure consists of colonies of alternating parallel layers of ferrite and cementite. Ferrite has a body centered cubic (bcc) structure with small amounts of carbon and substitutional elements in solid solution. The cementite is nearly pure Fe_3C. The interlamellar spacing varies between 157 and 209 μm. It was measured using the method of G.F. Vander Voort [10] and corresponds to a medium interlamellar spacing of an pearlitic steel with a mean hardness of 280 HV 10.

2.2   Results

Fatigue is an important mechanism of failure in rail steels. It is known that for pearlitic normalized steels, as they are used for rails, fatigue can already

Table 1. Chemical composition of the rail steel "UIC 60 900 A" [wt.-%]

C	Si	Mn	P	S	Cu	Cr	Ni	Mo	Co
0.677	0.234	1.054	0.022	0.009	0.131	0.108	0.041	0.009	0.005

Al, Ti, Sn, Mg, Ta, V, Nb, B, W $\leq$ 0.005

Table 2. Tensile properties of the rail steel "UIC 60 900 A"

Yield strength $R_{p0.2}$ [MPa]	Tensile strength R_m [MPa]	Elongation A [%]	Reduction in area Z [%]	Hardness [HV 10]
506	922	14	18	280

take place at loadings lower than the monotonic yield strength. It is also generally known that the cyclic properties depend on the loading conditions. This dependence is caused at the microscopic scale by typical changes in the dislocation structures [11,12]. Therefore it is necessary to have knowledge about the microstructure after defined cyclic loading. Typical examples of TEM micrographs of the unused rail steel are shown in Figure 1. Within the ferrite lamellae the dislocations are uniformly distributed and arranged in some kind of networks, except near to the cementite lamellae interface. The dislocation density is about $2.5 \cdot 10^9$. It seems to be relatively high for an unloaded state. The interface between the ferrite and cementite plates are very sharp. Some dislocations in the ferrite are arranged parallel to the interface (Fig. 1 (b)). This configuration was only found in the unused state.

a)   b)   c)
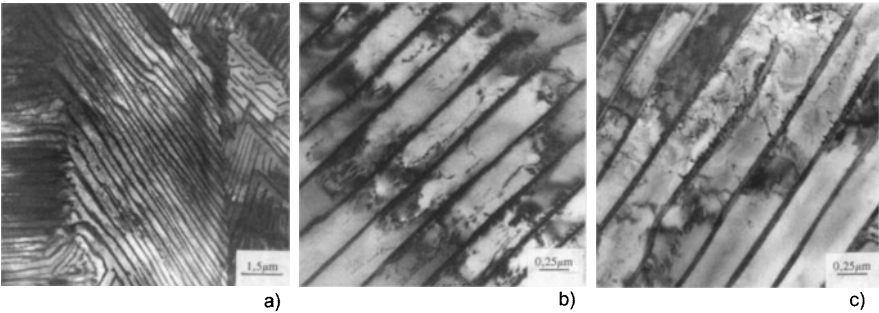
Fig. 1. (a) Overview. (b) Dislocations paralell to interface. (c) Dislocation-"networks"

Figures 2 show some typical dislocation structures of the specimen loaded $2 \cdot 10^6$ cycles without failure at $\varepsilon_{a,t} = 0.18$ %. On average the dislocation density is about $6 \cdot 10^9$ and thus approximately as high as in the unused condition. However, the interfaces between ferrite and cementite are no longer sharp. Only a few dislocation cells were found. Some dislocations inclined at an angle of about 45° to the cementite plates can be seen in Fig. 2 (c).

a)   b)   c)
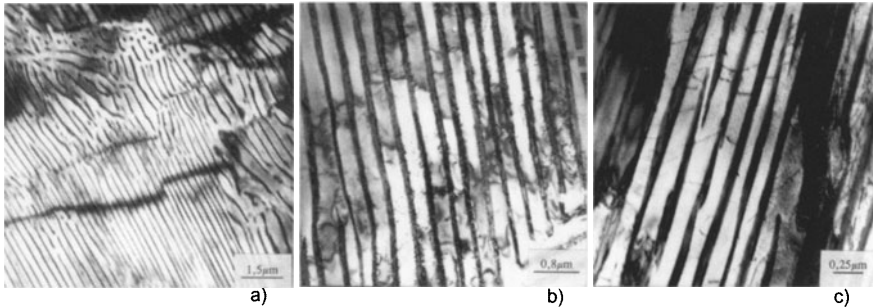
Fig. 2. (a) Overview. (b) Unsharp interfaces. (c) Inclined dislocations

In the specimen loaded at 1.0 % total strain amplitude ($N_f = 830$) a dislocation structure with a significant higher dislocation density is observed. As shown in Fig. 3 (a), (b) and (c) the dislocations are often arranged in a kind of cell structures. Dislocation walls located perpendicular to the cementite plates are forming the cells and determine the cell sizes. The cell walls are not packed very tightly (Fig. 3 (c)). The regions inside the cells seem to be largely free of dislocations. In other regions a diffuse dislocation structure (Fig. 3 (d)) with a significantly higher dislocation density than in the un- or low loaded states was oberved.

In samples taken from the used rail the dislocations are arranged in diffuse networks as well as in cells. Like in the highly stressed state the cells are formed by dislocation walls which are built up between cementite lamellae and wide regions free of dislocations between the walls (see Fig. 4 (a) - (c)). Obviously, the interfaces between ferrite and cementite which are sharp in the unused state (see Fig. 1) become somewhat diffuse during service loading. In addition some broken cementite lamellae were observed (Fig. 4 (d)).

2.3   Discussion

One reason of the conspicuous high dislocation density of the unloaded rail could be the roller straightening which is the last step in the production route of a rail [13]. There the cooled, distorted rail passes through a straightening machine with a series of rollers. While the rail is moving through this rollers,
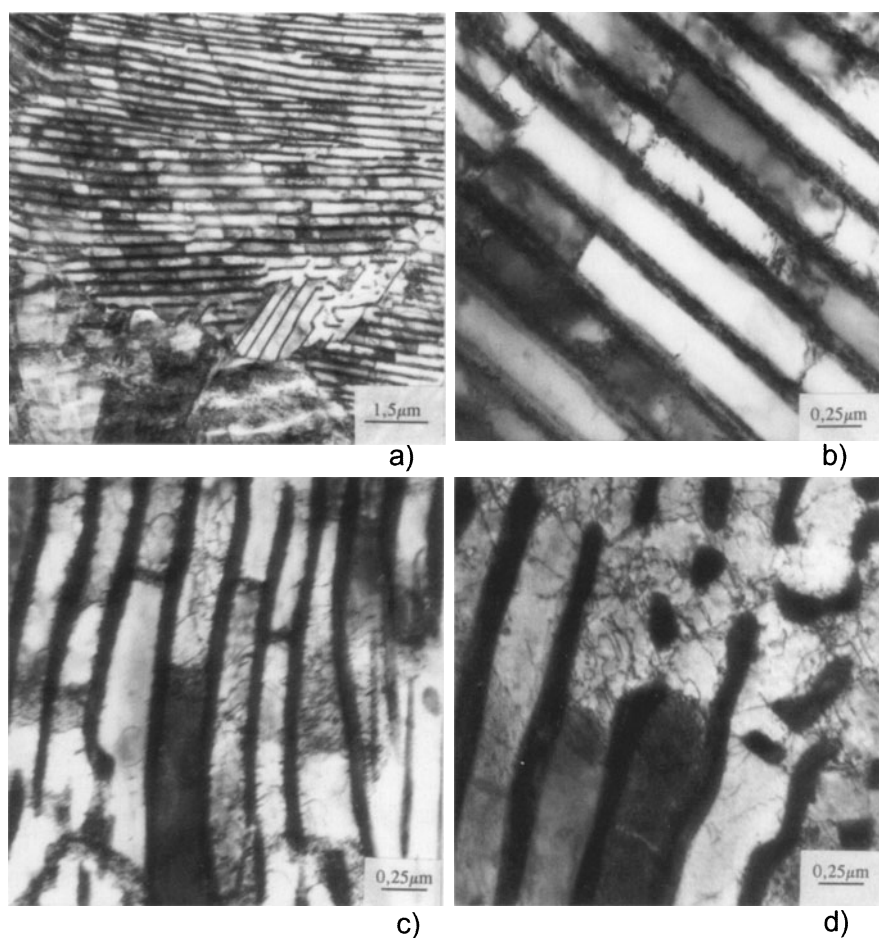
Fig. 3. (a) Overview. (b) Dislocations walls / cells. (c) Dislocations walls /cells. (d) Diffuse dislocation structure

which are located alternately below and above the rail, it is bent up and down and locally plastic deformation takes place [13]. The dislocation structure of the used rail is completely different to the one of the unused rail. In the unused state no dislocation walls and no broken cementite lamellae are found. But in the used state those structures have been found. Obviously the development of such structures is caused by the loading due to the accumulated traffic. This confirms to the observations of [14]. The blurred development of the ferrit-cementite interfaces is caused by the presence of dislocations near these interfaces. Therefore, it characterizes the grade of plastic deformation. For the unused rail the interfaces images are sharper than for the used rail. It is postulated that the dislocations are generated more easily in the soft
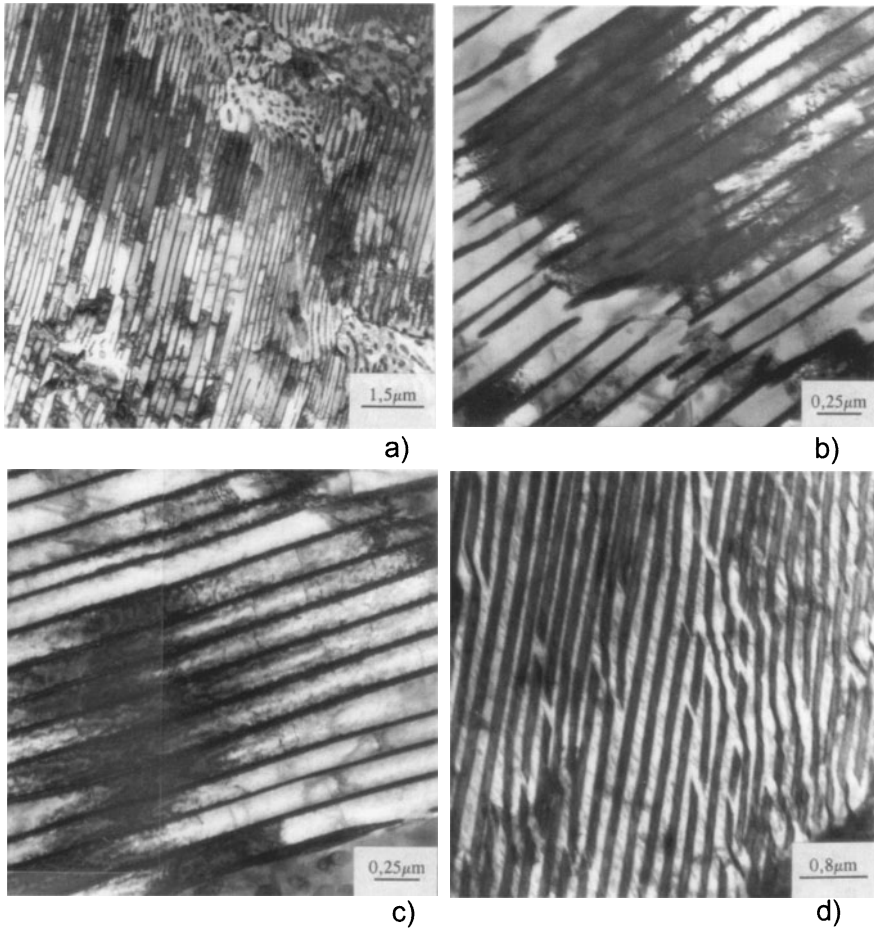
Fig. 4. (**a**) Overview. (**b**) Dislocation cells. (**c**)Dislocation cells. (**d**) Broken cementite lamellae

ferrite matrix than at the ferrite-cementite interfaces and that the dislocations along the interfaces are formed during cyclic deformation [9,14]. The driving force for cell formation is a certain amount of accumulated plastic strain [14]. Dislocations seem to move back and forth annihilating each other or setting down in an array. Both processes decrease energy to an amount lower than the energy state of a random dislocation distribution. This promotes cell wall formation, which is regarded as a thermally activated dynamic recovery process. The interlamellar spacing is expected to influence cell formation, because the free slip distance depends heavily on the spacing. As the interlamellar spacing increases cell formation becomes more likely. Obviously the stress fields which are combined with the pile up of dislocations

influences neighbouring ferrite lamellae across the cementite lamellae. So dislocation walls lying in one line over a couple of lamellae could be observed. It can be seen that increases in dislocation density and development of cell structure gradually becomes prominent with increasing cyclic loading. In the high loaded state the cell walls are packed more tightly and the regions in between the walls are almost free of single dislocations. These observations are also made by [9,14,15].

2.4   Conclusions

For a high and a low loaded rail steel a difference in dislocation structure was observed. In the high loaded state a typical structure consisting of cells formed by dislocation walls and dislocation free regions in between are detected in the ferritic lamellae. Those structures could not be found in the low loaded state. In the used rail frequently dislocation structures similar to the one of the high loaded specimen are observed. But there are also regions with diffuse dislocation networks which correspond to the low loaded state. Comparing the results of this investigation with the ones reported in literature, it is important to know that there is an influence of the specimen preparation, the used TEM microscope and the kind of measurement of the interlamellar spacing. Taking into account these influences the results reported here fit well to the results reported in the literature.

3   Deformation and Lifetime Behavior of a Cyclic Loaded Wheel Steel

For the last decades the accumulated service tonnage, the load of the rails and the train speed at Deutsche Bahn AG lines are increasing, causing some unexpected technical problems like corrugation, shelling and squats. To understand the complex phenomena of rolling contact fatigue in the rail/wheel system extensive investigations have to be done to describe the multiaxial rolling contact fatigue processes and to develop and verify multiaxial fatigue criteria and lifetime predictions. For these considerations a reliable fatigue data base of the utilized steels is required. To obtain such data, specimens were worked out of wheel rims used in high speed train wheels. With these specimens the lifetime behavior and the endurance limit are deduced from push-pull tests.

3.1   Material and Testing Specimens

The investigated wheel steel specimens were machined from the wheel rim of a monobloc wheel, grade R 7, manufactured by VSG Bochum. The heat treatment of the wheel is as followed: The whole wheel is austenitized and then only the wheel rim is quenched with water. Finally the whole wheel is

tempered at a minimum temperature of 500 °C. Unfortunately no more details about the heat treatment are available as these information are producer know-how. The wheel rim is a thick-walled component, so this heat treatment leads to a decreasing cooling gradient with increasing distance from the surface causing different microstructures within several regions of the wheel rim. Typical examples for the occuring microstructures are shown in Fig. 15. Near the surface up to a depth of about 2 mm a tempered martensite is found, which consists of a ferritic matrix including fine distributed cementite phase. In a depth of about 8 mm the microstructure contains ferrite, pearlite and bainite. In a depth of about 30 mm no more bainite is found and the microstructure consists of ferrite, pearlite and Widmannstätten structure. In deeper regions the Widmanstätten structure disappears and the grain size and the volume fraction of ferrite grows. Corresponding to the determined microstructure the hardness decreases from 290 HV 10 near the surface to 250 HV 10 in the middle of the wheel rim. As the microstructure changes continuously with distance from surface and also depends on the circumferential position of the wheel rim the various boundaries of the microstructure are difficult to locate. Solid cylindrical specimens were machined from the wheel rim with specimen axis parallel to the wheel axis, as shown in Fig. 16. The gauge sections contain the microstructures as follows: layer I: tempered martensite, layer II: ferrite, pearlite, bainite, layer III: ferrite, pearlite and Widmannstätten structure, layer IV: ferrite and pearlite. The specimens have a gauge length of 15 mm and a diameter of 7 mm in the gauge length. The results from measurements of the chemical composition and from tensile tests are given in Tables 3 and 4.

Table 3. Chemical composition of the wheel steel, grade R7 (wt.-%)

C	Si	Mn	Cr	Ni
0.541	0.268	0.772	0.238	0.149

3.2   Experimental Details

The specimens were loaded in push-pull tests on a Schenck 100 kN servo-hydraulic testing machine. Single-level stress controlled tests were performed in a loading range between 300 and 700 MPa. The loading cycle type was triangular. The specimens were tested at ambient temperature at a constant frequency (depending on the load) between 1 and 30 Hz. Testing was continued until either fracture occurred or $2 \cdot 10^6$ cycles without failure were reached.

Table 4. Mechanical properties of the wheel steel, grade R7

	Layer I	Layer II	Layer III	Layer IV
R_{eH} [MPa]	634	570	533	490
R_{eL} [MPa]	620	540	500	480
R_m [MPa]	880	866	857	836
A [%]	> 26	> 26	> 26	> 26
Z [%]	> 44	> 42	> 41	> 38

Pearlite and Ferrite

Pearlite, Ferrite and Widmannstätten Structure

Pearlite, Ferrite and Bainite

Tempered Martensite
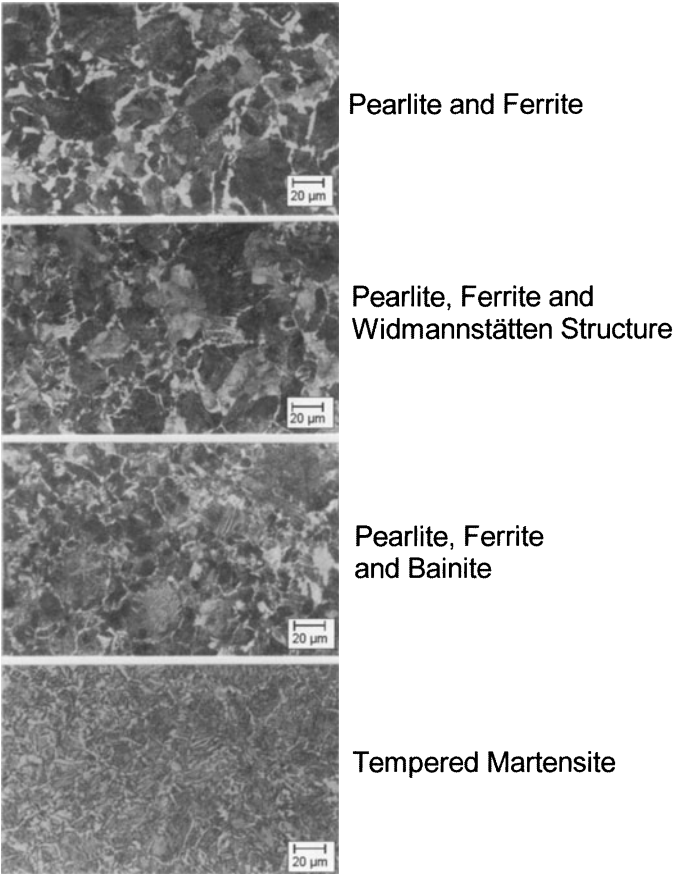
Fig. 5. Different microstructures in the wheel rim
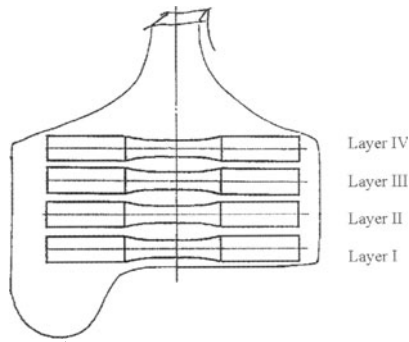
Fig. 6. Sampling of specimens

3.3   Results and Discussion

Cyclic Deformation Behavior. As the specimens are machined from different layers (indicated in Fig. 16) their cyclic deformation behavior corresponds to the different microstructures found in the gauge length. Figure 17 shows the deformation behavior of specimens taken from layer I. The chosen stress amplitudes yield characteristic behavior. With increasing stress amplitude the plastic deformation starts earlier, leads to higher plastic strain amplitudes and the lifetime decreases. The deformation behavior of specimens taken from layer III, see Fig. 18 at the same selected stress amplitudes, is quite different. Comparing the two selected layers it can be seen, that for similar stress amplitudes layer III is characterized by an earlier plastic deformation and larger plastic strain amplitudes. The cyclic deformation behavior of the several layers for a selected stress amplitude is shown in Fig. 19 and 19. The cyclic deformation behavior can be examined in the best way by a stress amplitude in low cycle fatigue range. The plastic strain amplitude $e_{a,p}$ versus number of cycles N, at 550 MPa stress amplitude, is plotted in Fig. 19. Layer I is characterized only by cyclic softening as it is known for a quenched and tempered microstructure [16- 19]. Curves of layers II,III and IV are similar, characterized by cyclic softening followed by cyclic hardening, typically for normalized ferritic-pearlitic microstructures [9,20]. Furthermore it can be seen that with increasing distance from surface cyclic softening starts earlier and leads to higher plastic strain amplitudes, due to the decreasing amount of bainite and the increasing content and grain size of ferrite. An increasing ferritic content combined with an increasing grain size of ferrite favors generation and movement of dislocations, enhancing cyclic plastic deformation.

The relationship between the mean strain $\varepsilon_{m,t}$ and the number of cycles N is shown in Fig. 19. There is also a good correlation to the microstructures in the different layers. The deeper the layer the earlier cyclic creep starts and higher mean strains are induced due to the same reasons as explained above.
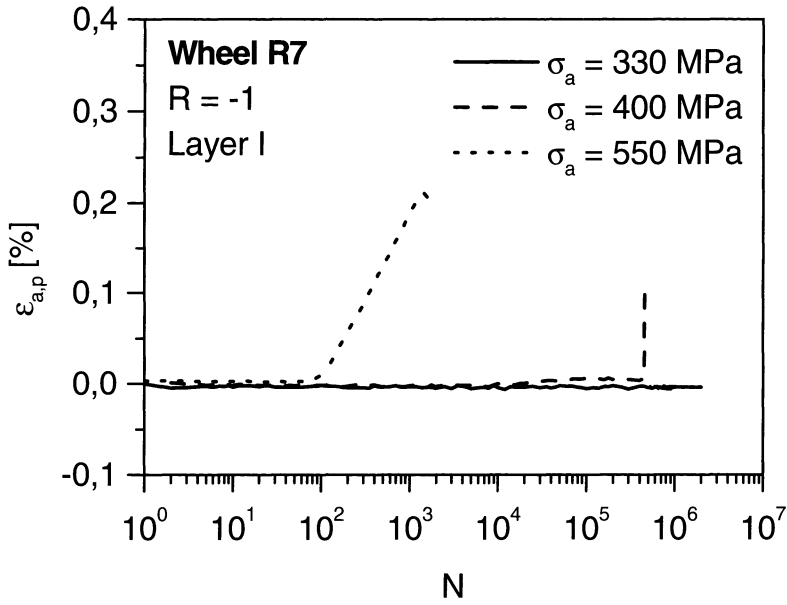
Fig. 7. Development of the plastic strain amplitude for specimens taken from layer I
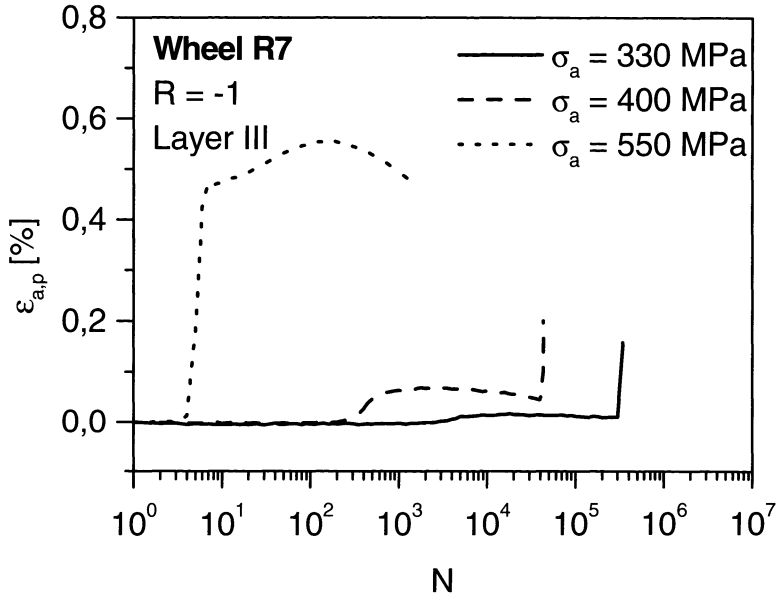
Fig. 8. Development of the plastic strain amplitude for specimens taken from layer II
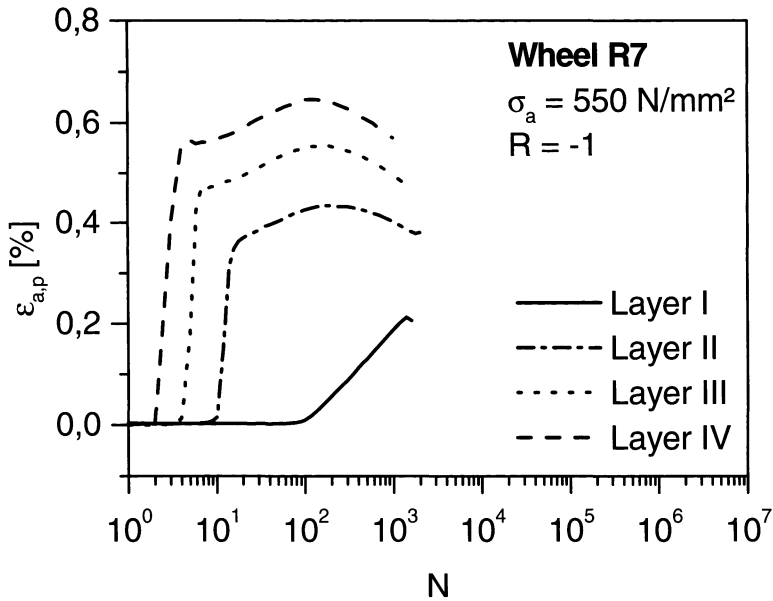
Fig. 9. Cyclic deformation behavior of the several layers
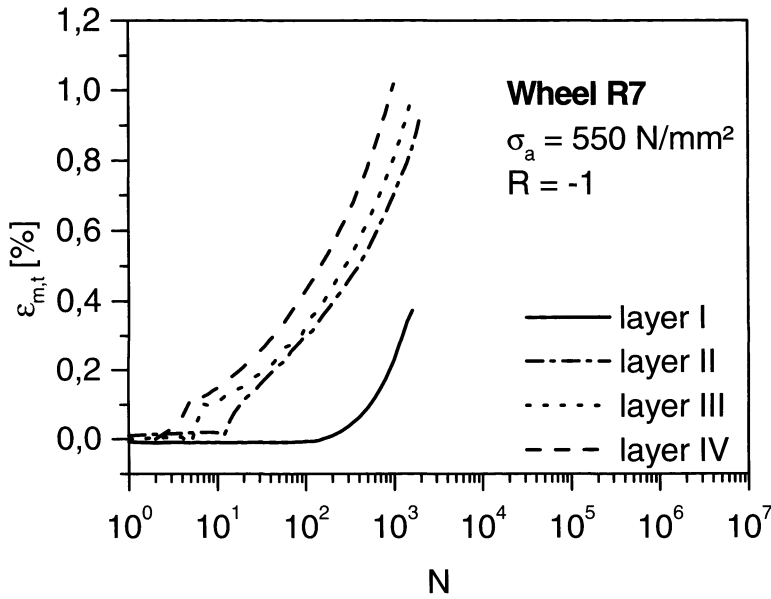
Fig. 10. Cyclic creep of the several layers

Lifetime Behavior. The different microstructures of different layers influence also the lifetime behavior as can be seen in Fig. 20. The endurance limit and the lifetime behavior in the high cycle fatigue range correspond to the determined cyclic deformation behavior. As with increasing distance from wheel surface plastic deformation and cyclic creep starts earlier and leads to higher mean strains and higher strain amplitudes the fatigue limit of these specimens decreases. But at higher stress amplitudes e.g. at $\sigma_a = 600$ MPa there is nearly no influence of the different microstructures on the lifetime behavior.
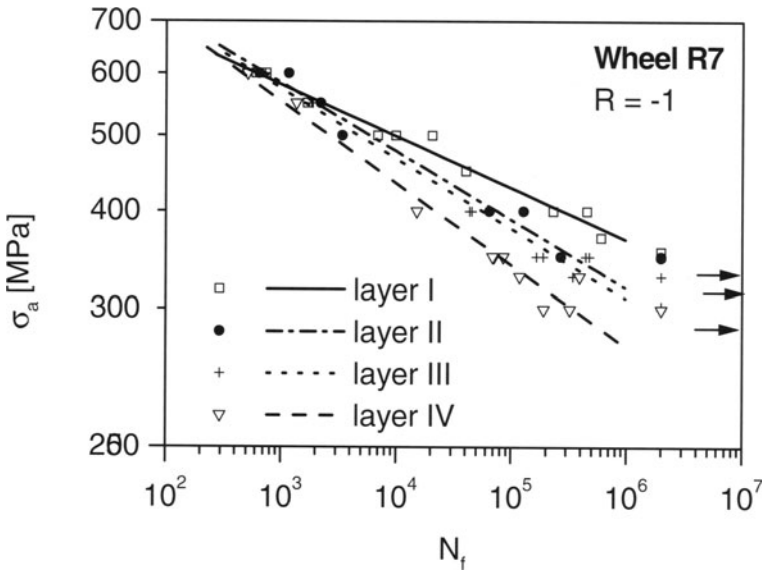
Fig. 11. Stress amplitude versus number of cycles to fracture of the several layers

3.4   Conclusions

Near the surface of a wheel rim a quenched and tempered microstructure is found changing continuously to a normalized microstructure at deeper regions. The specimens machined from the wheel rim contain these different microstructures. They influence the cyclic deformation behavior, the cyclic creep and the lifetime behavior. With increasing distance from surface the cyclic softening and cyclic creep starts earlier and leads to higher plastic strain amplitudes, and mean strains, while the fatigue limit decreases. Furthermore it was observed that cyclic deformation and lifetime behavior depends on the circumferential position. For similar distance from surface different cyclic

deformation behaviors have been observed, complicating systematic investigations. Obviously the actual heat treatment leads to an inhomogeneous and complex microstructure in the wheel rim. During long term service the wheel rim radius decreases as a consequence of wear and mechanical post processing, meaning that with increasing service time deeper layers, up to layer III, get in contact with the rail. All theses facts must be taken into account to understand the materials behavior under cyclic loading and to estimate fatigue life which should be the basis for a lifetime prediction.

References

1. B. Denne, K.-H. Lang, D. Löhe (1999) Mikrostruktur von Rädern und Schienen nach verschiedenen Belastungsfällen. In: R. Strauber, C. Lieser, R. Bütje, M. Bannasch (eds.) Werkstoffe für die Verkehrstechnik / Werkstoffwoche 98, Symposium 2, Wiley-VCH, Weinheim, ISBN 3-527-29939-4, 567-572
2. T. Beck, B. Denne, K.-H. Lang, D. Löhe (1999) Torsionsermüdungsprüfmaschine auf der Basis eines handelsüblichen AC-Servomotors. Werkstoffprüfung 1999, Bad Nauheim, DVM, Berlin, 291-300
3. B. Denne, K.-H. Lang, D.Löhe (2000) Deformation and Lifetime Behaviour of a Cyclically Loaded Wheel Steel. In: M. R. Bache et. all (eds.) Fatigue 2000 - "Fatigue & Durability Assessment of Materials, Components and Structures" Proc. Int. Conference FATIGUE 2000, 10-12 April 2000, Robinson College Cambridge, EMAS Ltd., West Midlands, UK, ISBN 1 901537 16 1 (2000), 387-394
4. B. Denne, K.-H. Lang, D.Löhe (2000) Deformation and Lifetime Behaviour of Cyclic Loaded Rail and Wheel Steels. In: P.J. Winkler et al. (eds.): Materials for Transportation Technology, EUROMAT 99-Volume 1, Wiley-VCH, Weinheim, ISBN 3-527-30124-0, 305-311
5. B. Denne, K.-H. Lang, D. Löhe (2000) TEM-Investigations of the microstructure of a cyclic loaded rail steel. Materials Week 2000, 25. - 28. September 2000, München, Poster B3-P1
   http://www.proceedings.materialsweek.org/proceed/mw2000_754.pdf
6. B. Denne, K.-H. Lang, D.Löhe (2001) Verformungs- und Lebensdauerverhalten des Schienenstahls "UIC 60 900 A". ZEV+DET Glasers Annalen - Die Eisenbahntechnik **125**, 2, 67-76
7. T. Beck, B. Denne, K.-H. Lang, D.Löhe (2001) Torsionsermüdungsprüfmaschine auf Basis eines handeslsüblichen AC-Servomotors. Materialprüfung, **43**, Heft 7-8, 283-287
8. H. Bühler, W. Schreiber (1957) Lösung einiger Aufgaben der Dauerschwingfestigkeit mit dem Treppenstufenverfahren. Archiv für das Eisenhüttenwesen, **28**, Heft 3, 153 - 156
9. D. Pilo (1979) Zum Wechselverformungsverhalten normalisierter unlegierter Stähle mit Kohlenstoffgehalten von 0,01 bis 1,02 Gew.-%. Dr.-Ing. Thesis, Universität Karlsruhe (TH), Germany
10. G.F. Vander Voort (1984) Measurement of the interlamellar spacing of pearlite. Metallography **17**, 1-17
11. M. Becker (1987) Das Wechselverformungsverhalten von CK 45 und CK 80 im Temperaturbereich zwischen 295 K und 873 K. Dr.-Ing. Thesis, Universität Karlsruhe (TH), Germany

12. D. Eifler (1985) Zusammenhang zwischen Mikrostruktur und Schwingfestigkeitsverhalten bei Stählen. In: D. Munz (ed.) Ermüdungsverhalten metallischer Werkstoffe, DGM-Verlag Oberursel, Germany, 7-9

13. G. Schleinzer (2000) Residual Stress Formation During the Roller Straightening of Rails", PhD-Thesis, Institut für Mechanik, Montanuniversität Leoben, Austria

14. H. Sunwoo (1982) Cyclic deformation of pearlitic eutectoid rail steel. Metallurgical Transactions A, **13A**, 2035-2047

15. P. Lukas, M. Klesnil (1992) Fatigue of metallic materials. Elsevier Scientific Publ. Comp., Amsterdam, The Netherlands

16. D.Eifler (1981) Inhomogene Deformationserscheinungen bei Schwingbeanspruchung eines unterschiedlich wärmebehandelten Stahles des Types 42 Cr Mo 4. Dr.-Ing. Thesis, Universität Karlsruhe (TH), Germany

17. V. Schulze (1993) Die Auswirkungen kugelgestrahlter Randschichten auf das quasistatische sowie ein- und zweistufige zyklische Verformungsverhalten von vergütetem 42CrMo4. Dr.-Ing. Thesis, Universität Karlsruhe (TH), Germany

18. D. Eifler, E. Macherauch (1983) Inhomogeneous work-softening during cyclic loading of SAE 4140 in different heat treated states. In: G.C. Sih, J.W. Provan (eds), Proc. Int. Symp. Defects, Fracture and Fatigue 2, Martinus Nijhoff/Dr. W. Junk Publ., The Hague, The Netherlands, 171-182

19. E. Macherauch, D. Eifler (1990) Microstructure and cyclic deformation of plain carbon and low alloyed steels. Int. J. Fatigue **12**, 165-174

20. D. Pilo, W. Reik, P. Mayr, E. Macherauch (1979) Cyclic Induced Creep of a Plain Carbon Steel at Room Temperature. Fatigue of Engineering Materials and Structures **1**, 287-295

Friction and Wear of Tractive Rolling Contacts

Ludger Deters, Steffen Engel, and Matthias Proksch

Otto-von-Guericke-Universität, Inst. für Maschinenkonstruktion, Lehrstuhl für Maschinenelemente und Tribologie, Universitätsplatz 2, 39106 Magdeburg

Abstract. Theoretical and experimental examinations of the tractive rolling contact are reported. The two roller contact is selected as the testing principle for friction and wear examinations to simulate the wheel-to-rail contact. The first part of the paper presents a computation model for describing the traction-creep behaviour under solid friction conditions. Methods of mathematical statistics are used to model the surface roughness. The deformation component of the friction force is calculated considering measured load-indentation depth curves. The adhesion component is determined by using the shear strength approach. Wear tests were performed on a new designed test machine. The wear volume rise turned out to be proportional to the increase of the acting pressure and the creep between the rollers. Raising the circumferential speed of the test rollers caused a reduction in the wear volume. A significant wear decrease at the driven rail rollers could be achieved by periodically reversing the direction of the acting friction forces. Examinations as to the extent to which theoretical results can be applied to practice, reveal a good correspondence.

1   Tasks and Objektives

Friction and wear of the wheel/rail system inflence the running performance of rail vehicles and the life of system components. A more realistic prevision about the wear behaviour of wheels and rails require further investigations into friction and wear behaviour within the wheel-to-rail contact. With a better understanding of the causes (friction) it will be possible to put the right construction on the effects (wear) and effectively counter damage to system components.

The present paper is aimed at developing a general computation model suitable for determining the traction coefficient between wheel and rail as well as its dependency on creep by using an appropriate model. In addition, this paper focussed on wear forecasts. To this end, the behaviour of the wheel and rail material will be experimentally analysed under various operating variables. Subsequently, the test data will be used to establish the wear-specific friction work $w_f = W_f/V_w$ from the quotient of friction work and wear volume. After the friction work $W_f = f_t\ F_n\ s_f$ with the traction coefficient f_t, the normal force F_n and the friction distance s_f has been calculated for the specific application, the wear volume V_w can be determined if the wear-specific friction work w_f is known.

Moreover, an attempt will be made to contribute to the application of results obtained from a model test bench to practice. To this end, wear results obtained from model tests will be compared with the values of a real wheel/rail system.

2   Two Roller Model

As an appropriate model for analysing friction and wear processes a roller/roller system has been chosen since such model is ideally suited to simulate kinematic and kinetic conditions occuring in a wheel-to-rail contact. In additon, predefined test parameters can be complied with. In order to ensure the comparability and the applicability of the results obtained from the two-roller test bench to the wheel-to-rail contact we employed identical materials, applied the same state and type of friction, observed a comparable contact ratio, used similar operating variables (pressure, creep, type of motion) and gave due consideration to identical wear mechanisms and forms of wear.

3   Calculation of Friction

3.1   Traction-to-Creep-Behaviour

The transmissible tangential force F_t and the traction coefficient $f_t = F_t / (f\ F_n)$ as a function of the creep $S = (r_1\ \omega_1 - r_2\ \omega_2) / (r_1\ \omega_1)$ with the roller radius r_1 and r_2 and the angular velocities of the rollers ω_1 and ω_2 can be calculated for the two roller contact by using Eq. (1) as suggested in [1] if the reduced radius $R^* = r_1\ r_2 / (r_1 + r_2)$ and half the contact width a as well as the friction coefficient f of the two roller contact are known (Fig.1). The friction coefficient f can be determined from the ratio between the maximum possible tangential force (friction force F_f) occuring under the condition of complete sliding in the contact and the normal load F_n.
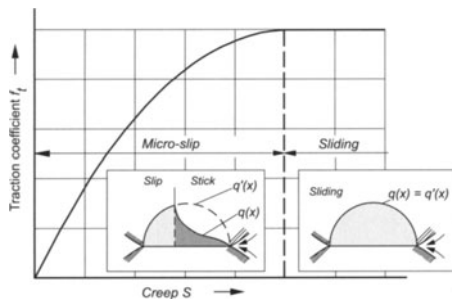
Fig. 1. Traction-creep curve

$$\frac{F_t}{f \, F_n} = 1 - \left(1 - \frac{S \, R^*}{a \, f}\right)^2 \qquad with \qquad f = \frac{F_f}{F_n} \tag{1}$$

If $S \, R^* / (a \, f) \geq 1$, $F_t = f \, F_n$ and sliding occurs everywhere in the contact. Otherwise and in addition to sliding, also sticking occurs at the leading area of the contact.

3.2   Description of the Tribotechnical System

The tribotechnical system under consideration consists of a wheel roller and a rail roller. The wheel roller is made of $R7$ steel, whereas the material of the rail roller is $900A$ steel. The diameters d_1 and d_2 and the width of the rollers are indicated in the chapter 4.2. The wheel roller and the rail roller run at different speeds, i.e. ω_1 and ω_2, and the rollers are subjected to various normal loads F_n. Apart from load, creep and external geometry, a number of surface roughness data and material parameters are of relevance to friction calculations.

Roughness Data

The mircogeometry is determined by performing two-dimensional measurements of the loaded surface of the rollers 1 and 2 at various points on their circumference in axial direction several times; to this end, a contact stylus intrument is used. Following [1], a simplified mathematical model comprising both rough surfaces can be used for describing the contact problem of rough surfaces, where a smooth rigid surface is moved relative to a resulting rough and deformable surface. To build the resulting surface, all measured roughness profiles of roller 1 and roller 2 are combined and summed up.
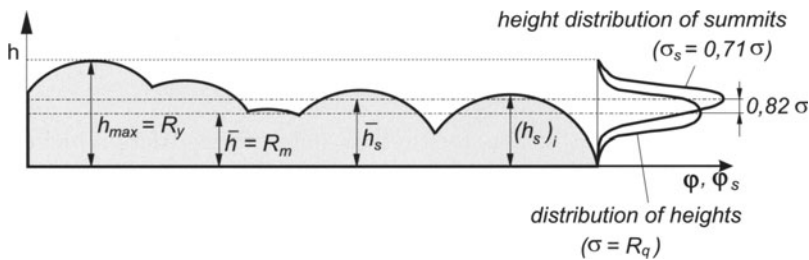
Fig. 2. Surface profile of the resulting model surface with amplitude density curves of heights and summits

The individual data sets of the resulting surface are for establishing the root mean squares R_q, the maximum profile valley depths R_m, the maximum profile heights R_y, the mean asperity radii $\overline{R_{as}}$, and the number of asperities

per unit length j_p^*, the latter corresponding to the number of evaluated radii per unit length. Subsequently, mean values are determined from all values of R_q, R_m, R_y, $\overline{R_{as}}$ and j_p^* and included in the computation model. The mean asperity radius $\overline{R_{as}}$ is determined by following the procedure as described in [2]. However, it is the distribution of asperity summits that is of relevance to friction calculations, rather than the height distribution of asperities (Fig.2). To establish the height distribution of asperity summits, the following assumptions are made corresponding to [1] or [5]:

- The heights of the asperity summits h_s are Gaussian-distributed with the standard deviation σ_s and mean value $\overline{h_s}$.This requirement is almost fulfilled by isotropic surfaces with a Gaussian height distrubution [1].
- Corresponding to [5] the standard deviation of the heights of asperity summits σ_s is described by the equation $\sigma_s = 0.71\sigma = 0.71\ R_q$ and the mean value of the summit heights $\overline{h_s}$ is given by $\overline{h_s} = \overline{h} + 0.82\sigma = R_m + 0.82R_q$.
- The value of the maximum profile height R_y of the asperity summits is identical to that of the resulting model surface.
- The asperity summits exhibit a constant mean radius, which is identical to the mean radius $\overline{R_{as}}$ of the asperities from the resulting surface.

The standardised distribution density function of the asperity summits $\phi_s(\xi)$ on the resulting model surface satisfies the equation [4]:

$$\phi_s(\xi) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\xi - \overline{\xi}\right)^2\right] \quad with \quad \xi = \frac{h_s}{\sigma_s} \quad and \quad \overline{\xi} = \frac{\overline{h_s}}{\sigma_s} \quad (2)$$

The asperities j_a on the nominal contact area A_a can be calculated by means of the number of asperities per area j_a^* with $j_a = j_a^*\ A_a$. The number of asperities per area is estimated with $j_a^* = 1.21\ j_p^{*\,2}$ following the random field theory as suggested by [6].

Material Data

The material data of the test rollers required for determining friction include the hardness HV 0.1 of roller 1 and 2, the moduli of elasticity E_1 and E_2, and the Poisson's ratios ν_1 and ν_2. These data can be used for determining the reduced modulus of elasticity $E^* = 2E_1E_2/\left[E_1\left(1 - \nu_2^2\right) + E_2\left(1 - \nu_1^2\right)\right]$ and the yield strength R_e of roller 1 and roller 2 on the basis of the relation $R_e \approx HV0.1/c_y$ with $c_y = 3$.

3.3   Contact Parameters under Tribological Loading

Compared to smooth surfaces, rough surafces exhibit an increased contact width and, hence, a larger nominal contact area. The following relation holds

true: $A_{a\ rough} = A_{a\ smooth}\ a^*$, where $a^* = a_{rough}/a_{smooth}$ is the effective contact half-width and $A_{a\ smooth}$ is the nominal Hertzian contact area for smooth surfaces. The effective contact half-width depends on the operating variables, marcogeometry, surface roughness and the material properities,as suggested by [7]. The tests under consideration yielded a mean value of $a^* \approx 1,24$. In the range $0 \le z \le z_{max}$ the number of asperity contacts j_r can be calculated as follows (Fig.4):

$$j_r = j_a \int_z^{z_{max}} \phi_s\left(\xi\right) d\xi \quad with \quad z = \frac{h}{\sigma_s} \quad and \quad z_{max} = \frac{h_{max}}{\sigma_s} \qquad (3)$$

To establish the load-carrying capacity F_n of a surface as a function of the approximation δ, indentation tests with a diamond Brinell ball ($400\mu m\ dia.$) were performed. Both the loading and the indentation depth were recorded. Standardising the test load with the normal load $F_{n,y}$ under which initial plastic deformations occur, and the depths of indentation with the aproximation δ_y under which initial plastic deformations occur, yields the standardised load-indentation depth curve as shown in Fig. 3. The values of δ_y and $F_{n,y}$
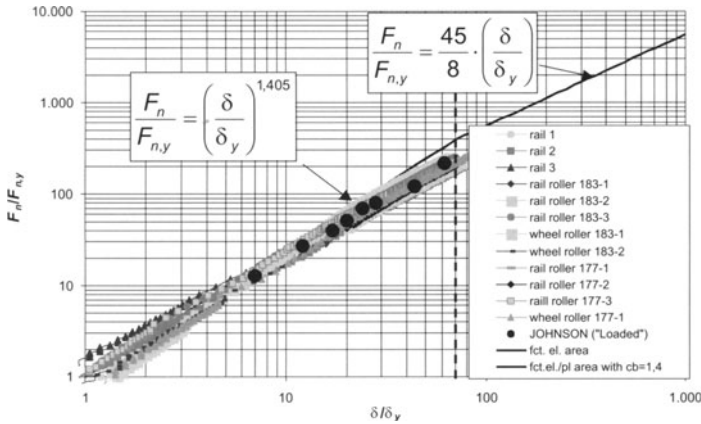
Fig. 3. Approximation of standardised load-indentation depth curves

can be determinend by using the following equations:

$$\delta_y = 25.27\overline{R_{as}}\left(\frac{R_e}{E^*}\right)^2 \quad and \quad F_{n,y} = 84.67R_e^3\left(\frac{\overline{R_{as}}}{E^*}\right)^2 \qquad (4)$$

Three areas can be distinguished in the standardised load-indentation depth curve: (i) elastic deformation, (ii) elastic/contained plastic deformation, and (iii) plastic deformation. These three areas can be described by using an approximation function, i.e. $F_n/F_{n,y} = (\delta/\delta_y)^{3/2}$ for the elastic area, $F_n/F_{n,y} =$

$(\delta/\delta_y)^{7/5}$ for the elastic/contatined platstic area, and $F_n/F_{n,y} = 5.63\ (\delta/\delta_y)$ for the plastic area. Calculations of the surface approximation u are based on the following assumptions: A smooth rigid surface approximates to the resulting rough deformable model surface (Fig. 4). When the smooth rigid surface approaches the resulting model surface, the asperities on the model surface are deformed. Depending on the extent of approximation, deformations are either elastic, elastic/contained plastic (cavity model [1]) or plastic.The variables in Fig. 4 are defined as follows:

$$z_y = \frac{\delta_y}{\sigma_s}; \quad z_\delta = \frac{\delta}{\sigma_s}; \quad \delta = (\xi - z)\,\sigma_s$$
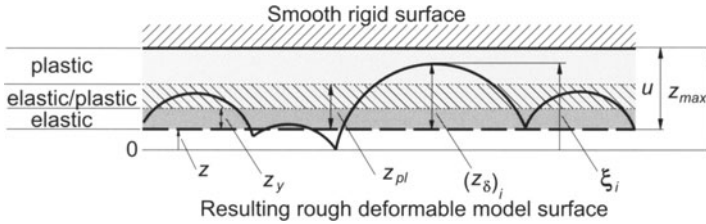
Fig. 4. Possible deformation areas when the smooth rigid surface approximates to the rough surface (bolt dashed line marks the value u by which the smooth rigid surface approaches the rough surface)

From the equation below for calculating the load-carrying capacity F_n of the asperities, the distance z between the zero line of the resulting model surface and the approximated rigid smooth surface can be determined iteratively until an equilibrium of forces is reached. The load-carrying capacity F_n of the resulting model surface can be determined, e.g. for the plastic range $(0 \leq z \leq z_{max} - z_{pl})$, by using the following equation:

$$F_n = \frac{2}{3}E^*\sqrt{\overline{R_{as}}}\ j_a\sqrt{\sigma_s^3}\int_z^{z+z_y}\phi_s\,(\xi)\sqrt{(\xi - z)^3}\,d\xi\ + \tag{5}$$

$$+\ F_{n,y}\ j_a\left(\frac{\sigma_s}{\delta_y}\right)^{c_b}\int_{z+z_y}^{z+z_{pl}}\phi_s\,(\xi)\,(\xi - z)^{c_b}\,d\xi +$$

$$+\ 2c_y\pi R_e\overline{R_{as}}\ j_a\sigma_s\int_{z+z_{pl}}^{z_{max}}\phi_s\,(\xi)\,(\xi - z)\,d\xi$$

The load-carrying capacities of the asperities in the elastic and the elastic/contained plastic range can be calculated similary to Eq. (5) by leaving out those terms, which are no longer required, and adjusting the integration limits. If z is determined (iteratively) for the given external load, the real

contact area A_r can be calculated, e.g. for the plastic range $(0 \leq z \leq z_{max} - z_{pl})$, by using the following equation:

$$A_r = \pi \overline{R_{as}} \, j_a \sigma_s \int_z^{z+z_y} \phi_s \left(\xi \right) \left(\xi - z \right) d\xi + \tag{6}$$

$$+ \ m\pi \overline{R_{as}} j_a \sigma_s \int_{z+z_y}^{z+z_{pl}} \phi_s \left(\xi \right) \left(\xi - z \right) d\xi \ +$$

$$+ \ (1 - m) \, \pi \overline{R_{as}} \, j_a \delta_y \int_{z+z_y}^{z+z_{pl}} \phi_s \left(\xi \right) d\xi \ +$$

$$+ \ 2\pi \overline{R_{as}} \, j_a \sigma_s \int_{z+z_{pl}}^{z_{max}} \phi_s \left(\xi \right) \left(\xi - z \right) d\xi$$

3.4   Deformation Component of Friction Force

The deformation component $F_{f,def}$ of the friction force F_f is calculated for two asperities, which are typical for the individual surfaces and exhibit a radius $\overline{R_{as1}}$ and $\overline{R_{as2}}$, respectively. These asperities are moved relative to each other, approach and deform each other (Fig. 5). This approach does not facilitate any detailed examination of individual microcontacts, but provides an integral value for the deformation component of the friction force.

Fig. 5. Contact of two asperities typical for the surfaces

The representative asperities in Fig. 5 are exposed to a mean normal load $F_{n,r} = A_r/j_r$. The real contact area of the representative asperities in a centric contact is assumed to be $A_{r,r} = A_r/j_r$. As a result, the mean pressure $p_m = F_n/A_r$ occurs in the representative mircocontact.

When the representative asperities of body 1 and body 2 move relative to each other, both asperities together are flattened by δ_r and the deformation

work $W_{f,def,r}$ is done. It is assumed that this deformation work $W_{f,def,r}$ corresponds to the friction work $W_{f,def,r}$, which (as a result of deformation) is carried out when the representative asperities move relative to each other. The friction work $W_{f,def,r}$ results from the component $F_{f,def,r}$ acting along the friction distance s_f, multiplied by the friction distance s_f.

$$W_{def,r} = W_{f,def,r} = F_{f,def,r} \, s_f \tag{7}$$

Assuming that the major part of the deformation work $W_{def,r}$ results from deforming the representative asperities in normal direction (y-direction in Fig. 5), only deformations resulting from normal loads are considered for calculating the friction component $F_{f,def,r}$.

$$W_{def,r} = \int F_{n,r} \left(\delta_r \right) d\delta_r \tag{8}$$

The deformation work $W_{def,r}$ is determined taking the indentation curve (Fig. 3) into account. After establishing whether the deformation is elastic $(F_{n,r} \le F_{n,y})$ or elastic/contained plastic $(F_{n,y} \le F_{n,r} \le F_{n,pl})$ or plastic $(F_{n,pl} \le F_{n,r})$, the approximation δ_r of the representative asperities can be derived from the contact area $A_{r,r}$ of the representative asperities in the centric contact as follows: For elastic deformation δ_r is calculated from $\delta_r = A_{r,r} / \left(\pi \overline{R_{as}} \right)$, for elastic/contained plastic deformation from $\delta_r = \left(A_{r,r} - (1-m) \pi \overline{R_{as}} \delta_y \right) / \left(m \pi \overline{R_{as}} \right)$ with $m \approx 2$, and for plastic deformation from $\delta_r = A_{r,r} / \left(2 \pi \overline{R_{as}} \right)$.

When δ_r is known, the deformation work $W_{def,r}$ at the representative asperities results from the solution of Eq. (8) using the approximation functions from the standardised load-indentation depth curves for $F_{n,r}$. The elastic component can be neglected.

Theoretical examinations in [12] show that for a radius ratio of $\overline{R_{as1}} / \overline{R_{as2}} = 0.5 \ldots 2$ the friction distance s_f is almost identical with the diameter $d_{r,r}$ of the contact area of the representative asperities $A_{r,r}$ (Fig. 5). If the deformation work $W_{def,r}$ is known, this means for calculating the deformation-induced friction force $F_{f,def,r}$ at the representative individual contact that:

$$F_{f,def,r} = \frac{W_{def,r}}{d_{r,r}} \quad with \quad d_{r,r} = \sqrt{\frac{4A_{r,r}}{\pi}} \tag{9}$$

The total deformation-induced friction force $F_{f,def}$ can be determined by using the number of contacts j_r for the following equation:

$$F_{f,def} = j_r \, F_{f,def,r} \tag{10}$$

3.5   Adhesion Components of Friction Force

The effect of adhesion in the friction process (adhesion component $F_{f,ad}$ of the friction force F_f) results from the seperation of adhesive bondings in the real contact area. Compared to other models, it is fairly easy to determine all input variables required for the shear strength model, and hence the adhesion component $F_{f,ad}$ of the friction force F_f is calculated by following the approaches of [13]:

$$F_{f,ad} = \tau_{elpl} \ A_{r,ad,elpl} + \tau_{pl} \ A_{r,ad,pl} \quad , \tag{11}$$

where τ_{elpl} and τ_{pl} are the shear strengths of the adhesive bondings at elastically/contained plastically and plastically deformed asperities, respectively. $A_{r,ad,elpl}$ and $A_{r,ad,pl}$ are the real contact areas of the elastically/contained plastically and plastically deformed asperities, respectively. Adhesion is only to occur in these contact areas as a deformation with plastic components causes the formation of fairly large real contact areas and the breaking-up of existing absorption and oxide layers, intensifying molecular interactions between the surfaces. The shear strengths τ_{elpl} and τ_{pl} in Eq. (11) are calculated as follows: $\tau_{elpl} = m_{elpl} \ \tau_s$ and $\tau_{pl} = m_{pl} \ \tau_s$, where τ_s is the shear strength of the softer material. Due to the presence of oxide layers and other contaminants in the contact area, it is assumed that the shear strenth τ_s of adhesive bondings is smaller than that of pure materials. This aspect is given consideration by $m_{pl} = 0.8 \ldots 0.95$ and $m_{elpl} = 0.6 \ldots 0.85$. Under the condition that shakedown of the boundary subsurface layers is almost completed or does not occur any longer, the shear strength can be determined by using the relation $\tau_s = HV0.1/ \ 6$. If plastic deformations occur at the asperities $(0 \leq z \leq z_{max} - z_{pl})$, the real ('adhesion-effective') contact area $A_{r,ad,elpl}$ can be obtained by using the following equation:

$$A_{r,ad,elpl} = m\pi\overline{R_{as}}j_a\sigma_s \int_{z+z_y}^{z+z_{pl}} \phi_s\left(\xi\right)\left(\xi - z\right)d\xi \ + \tag{12}$$

$$+ \left(1 - m\right)\pi\overline{R_{as}} \ j_a\delta_y \int_{z+z_y}^{z+z_{pl}} \phi_s\left(\xi\right)d\xi$$

If also shearing τ occurs in the contact area, apart from the pressure p, the contact area $A_{r,pl,0}$ occurring in a plastic contact under a mere normal load F_n, increases to $A_{r,ad,pl}$. In the relevant literature this phenomenon is called 'junction growth'. The following holds true:

$$A_{r,ad,pl} = \sqrt{A_{r,pl,0}^2 \left[1 + \alpha\left(\frac{F_t}{F_n}\right)^2\right]} \tag{13}$$

A value of $\alpha = 9$ was chosen for the calculations. The real contact area $A_{r,pl,0}$ at plastically deformed asperities in the initial state (without the effect of tangential forces) corresponds to that part of the real contact area where the asperities are plastically deformed under the influence of a mere normal load. Within the range $0 \leq z \leq z_{max} - z_{pl}$ (plastic deformation):

$$A_{r,pl,0} = 2\pi \overline{R_{as}} \, j_a \sigma_s \int_{z+z_{pl}}^{z_{max}} \phi_s\left(\xi\right)\left(\xi - z\right) d\xi \tag{14}$$

Calculation of the increased real ('adhesion-effective') contact area $A_{r,ad,pl}$ according to Eq. (13) requires an iterative approach as the ratio F_t/F_n is not known in the beginning.

3.6   Friction Coefficient

The total fricton force F_f consists of two components, i.e. $F_{f,def}$ and $F_{f,ad}$. Hence,the friction coefficient f of the two roller contact can be calculated as follows:

$$f = \frac{F_{f,def} + F_{f,ad}}{F_n} \tag{15}$$

Following the determination of f, the transmissible tangential force F_t or the traction coefficient f_t can be calculated.

4   Test Machine and Test Rollers

4.1   Test Machine Design

A new two roller test machine was built to simulate the wheel-to-rail contact and examine friction and wear behaviour. The test machine was provided with a novel loading system for applying the load to the test rollers symmetrically from two sides by means of supporting rollers. This design virtually prevents the test rollers from misaligning (Fig. 6). With regard to friction and wear a high reproducibility of test results was obtained.

4.2   Test Rollers

The test rollers were machined from original components (wheel and rail) to simulate the wheel/rail system under most realistic conditions, if possible. Hence,certain restrictions had to be observerd as to their dimensions. Fig. 7 depicts the points from where the standard test rollers were taken as well as their dimensions.
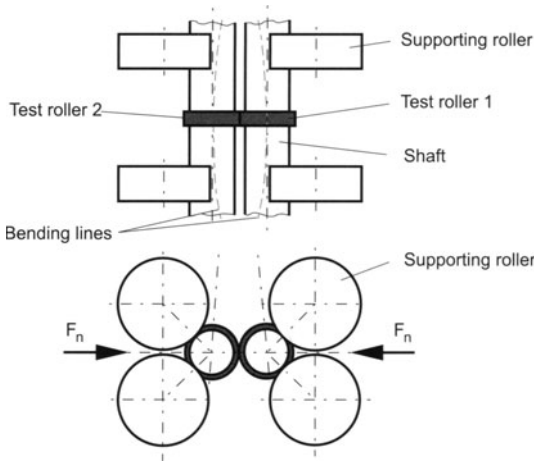
Fig. 6. Loading principle of the two roller test machine
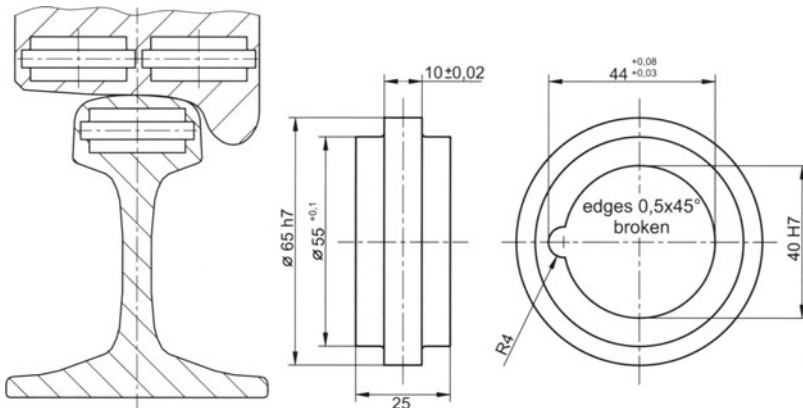
Fig. 7. Points from where test rollers were taken (left) and dimensions of test rollers (right)

5   Results

5.1   Comparsion Between Theoretical and Experimental Traction Coefficients

Fig. 8 presents a comparison between the measured and the calculated traction-to-creep curves (test no. 6 : 0 to 2.5% creep; test no. 7 : 0 to 4 % creep). A comparsion between the traction coefficients measured at the two roller test machine and those calculated reveals a good correspondence. It is assumed that deviations are caused by roughening of the test rollers, varying more markedly at higher creep values. Hence, uncertainties are encountered when determining roughness values. Strikingly, the curves exhibit a rather flat

pattern. This might be attributable to the fact that the maximum traction as a function of creep and pressure differs at the various operating points due to changes in the mircogeometry and the material properies in the subsurface layer.
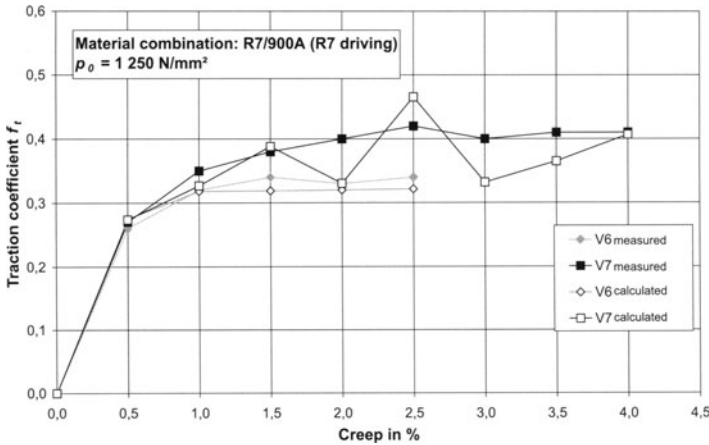
Fig. 8. Comparsion between measured and calculated traction coefficients

5.2   Experimental Friction and Wear Data

Influence of Pressure

Experimental examinations revealed a slight decrease in the measured traction coefficients when the pressure between the two test rollers increased (Fig. 9). However, at both the driving roller and the driven roller the pressure rise causes a continuous growth of the wear volume. As a result of the traction coefficient decreasing slightly and wear volume increasing more intensely, the wear-specific friction work follows a decreasing path (Fig. 10).This indicates the lower wear resistance of the tested material combination when the pressure grows higher.

For the driving roller this wear behaviour is probably due to a slightly increasing oxide film and a splitting off from the surface at higher pressures. This spalling away of oxide particles exposes more metall which rapidly oxidizes once again. At the driven roller the predominant metal surface is stressed higher which leads to an increased production of lamellar wear particles.

Influence of Circumferential Speed

When the circumferential speeds of the test rollers are increased while keeping the creep and the number of contacts ($N_c = 200,000$) constant, the traction
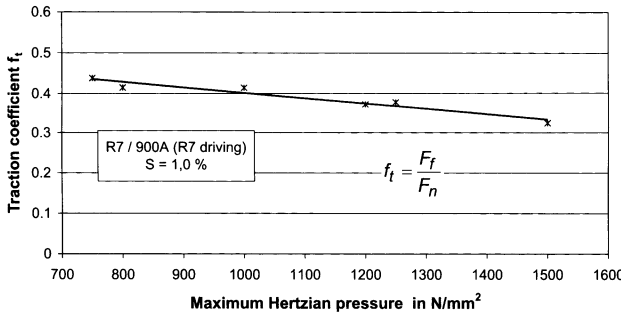
Fig. 9. Traction coefficients as a function of maximum Hertzian pressure
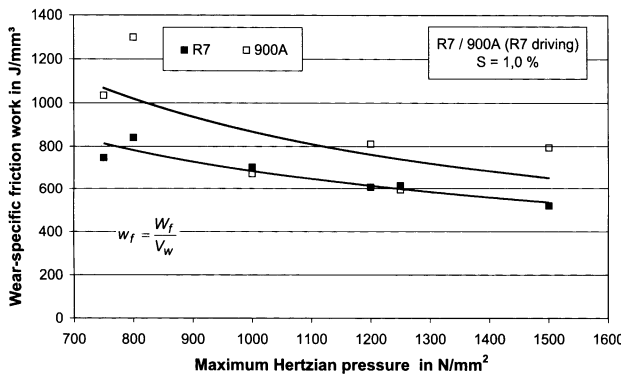
Fig. 10. Wear-specific friction work under varying pressure conditions
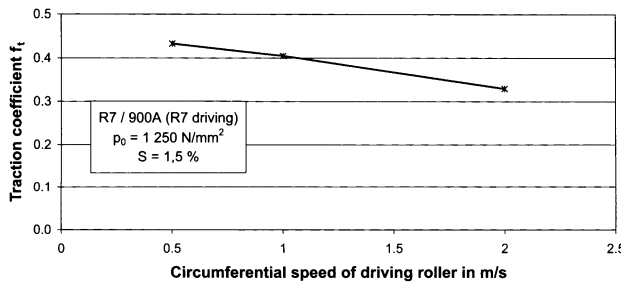
Fig. 11. Traction coefficients at various circumferential speeds

coefficients slightly decrease (Fig. 11). Also the wear volume after 200,000 revolutions each of the driving roller is reduced when the speeds of the rollers are increased.

The lower wear volume and the resulting higher wear resistance at higher circumferential speeds, as presented in Fig. 12, might be attributable for the driving roller to a more continuous and thicker oxide film which becomes possibly weaker as a result of the growing film thickness and the increasing

temperature at rising velocities. For the driving roller the shorter contact times at higher speeds cause higher deformation speeds in the subsurface layer of the material. Higher deformation speeds may exert a favourable influence on the strength of the material and, hence, on wear resistance, at least in the speed range under consideration.
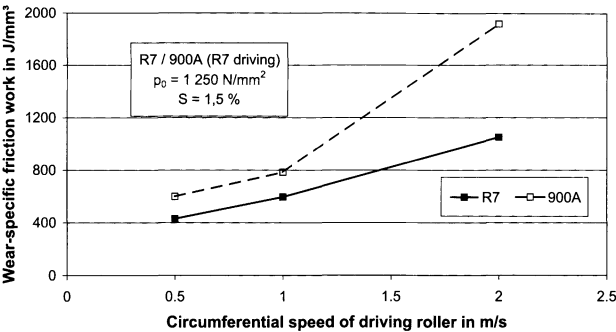
Fig. 12. Wear-specific friction work at varying circumferential speeds

Influence of Creep

In the tests involving a change in creep, the dry rolling/sliding contact exhibited the known traction-to-creep behaviour(Fig. 13). The traction behaviour changes only slighty when the rail roller is used as the driving roller, corresponding to brake processes in the wheel/rail system and marked by negative creep values in the diagrams.
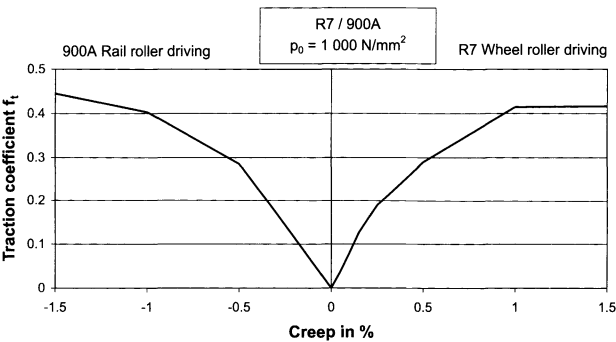
Fig. 13. Measured traction-to-creep curve (left: rail roller driving, right: wheel roller driving)

As can be cleary seen in Fig. 14 for negative creep values, there is a good correlation between wear volume development and hardness values obtained

on the tribological loaded surface. In the tests involving lower creep values
(≤ 0.25 %) both rollers were completely covered with a reddish-brown and
shiny oxide layer. Due to the resulting protective effect there was hardly
any wear. Wear did not clearly increase until after the creep values rose,
causing a destruction of the oxide layers as a result of higher shear stresses.
These changing wear mechanisms, i.e. from a mere tribochemical reaction
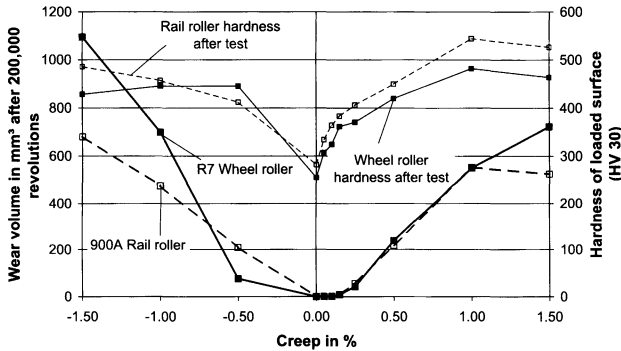
Fig. 14. Wear volume and hardness of the running tread as a function of creep
(left: rail roller driving, right: wheel roller driving)

at both rollers to the fatigue and abrasion mechanisms at the driven roller
and tribochemical reaction and fatigue (of the reaction layers) at the driving
roller are also reflected by the values of wear-specific friction work (refer
to Fig. 15), i.e. a change was observed from low wear levels (high wear-
specific friction work values) to high wear levels (low wear-specific friction
work values) as had been described earlier, for instance, in [14]. If the creep
exceeds a value of about 0.25 % ,i.e. the critical value for the wear of this load
condition $\left(p_o = 1,000\ N/mm^2\right)$ marking the transition from lower to higher
wear levels, wear resistance takes on almost constant values.

Influence of Reversing Directions of Loading

Tests were performed to simulate travel motions on a single-track line where
the rails are loaded alternately in both directions (refer to Fig.16). The rolling
direction of both rollers on the test machine was changed at intervals (e.g.
after 20,000 revolutions each of the driving roller). As can be demonstrated
by a comparison between the traction coefficients under reversing and con-
stant loads, this constant alternation of the loading directions has hardly any
effect on the transmissible tangential forces (reversing directions of loading:
$f_t = 0, 4$; constant direction of loading: $f_t = 0, 41$). However, a significant dif-
ference occurs after 200,000 revolutions of the driving wheel roller regarding
the wear volume of the rail roller, i.e. the wear volume decreases almost by
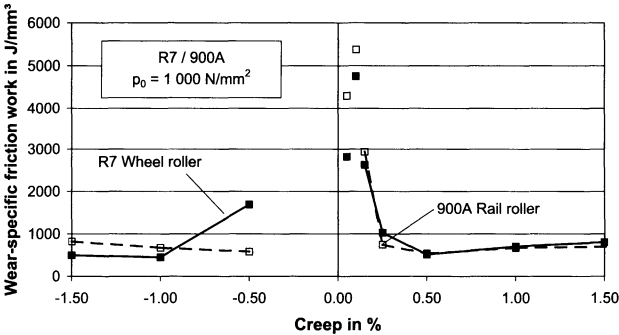
Fig. 15. Wear-specific friction work at varying creep values(left: rail roller driving, right: wheel roller driving)
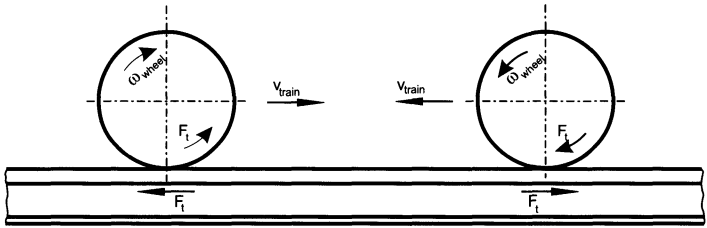
Fig. 16. Travel motions and tangential forces of a shuttle train on a single-track line

half. In contrast to that, hardly any difference can be observed as to the wear of the wheel roller. This varying wear behaviour of the rail roller under reversing and constant directions of loading is also reflected by the wear-specific friction work (refer to Fig.17). The different wear behaviour under reversing directions of loading is due to structural changes in the subsurface layer of the wheel roller and the rail roller.
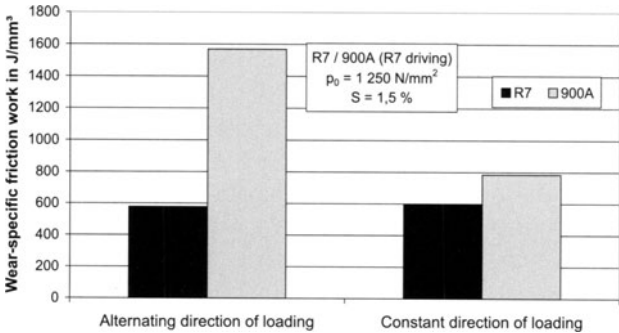
Fig. 17. Wear-specific friction work under reversing directions of loading

5.3   Comparison Between Wear Occurring in Original Rails and Test Rollers

A comparsion with wear values obtained from track measurements was performed to check the wear values obtained on the test bench. To this end, the wear-specific friction work was determined for selected DB AG track sections. Eq. (16) was used for calculating the wear-specific friction work of the original rail. The values (f_t, F_n, s_f, h_w, A_a) required for the calculations were obtained (i) from data of the original system, e.g. the wear height $h_{w,r}$ and the number of roll-over cycles required for determining the sliding distance s_f, and (ii) from analyses of the wheel-to-rail contact as published in the literature, e.g. [11], [16] and [17]. The wear height $h_{w,r}$ was determined by profile measurements; for calculation purposes the wear value measured in the middle of the rail head was used. The number of roll-over cycles during the measuring period can be obtained from the period between the measurements and the daily line load with the help of the mean axle load occuring on the examined DB AG track. Presently, the input variables for the mean pressure $\overline{p} = F_n/A_a$, the traction coefficient f_t and the nominal area of the contact A_a need to be determined from data published in the literature.

$$w_{f,r} = \frac{\alpha_s w_f}{V_{w,r}} = \frac{\alpha_s\ f_t\ \overline{p}\ s_f}{h_{w,r}} \qquad with\ \ \alpha_s \approx 0.5 \tag{16}$$

For the original system a range is indicated of wear-specific friction work that is likely to occur as there are no defined operating conditions for the overall measuring period. The possible range of contact conditions is limited by the two extreme cases of fully lubricated and dry contact conditions. The exact wear-specific friction work values of the wheel/rail system range between these two limit values. Fig. 18 depicts the range of the original system marked grey.
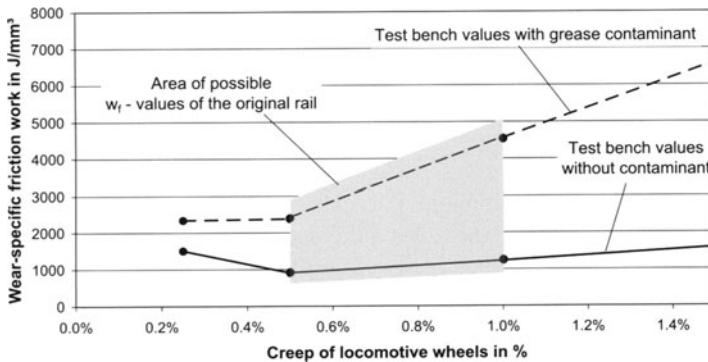
Fig. 18. Comparison between wear-specific friction work values obtained on the test bench and values likely to occur in practice

The values obtained from the model tests are in fairly good correspondence with the area of the original system. This good correspondence shows that results obtained from model test can be applied to railway practice when relevant conditions are fulfilled.

6   Summary

Theoretical and experimental examinations of the rolling contact were reported. The two roller contact was selected as the testing principle by using the standard-quality materials R7 for the wheel roller and 900A for the rail roller. The maximum Hertzian pressure and the creep were varied.

The first part of the paper presents a computation model for describing the traction-creep behaviour in the rolling contact under solid friction conditions. Calculations are divided into two major sections (i) calculation of surface approximation and of the real contact area and (ii) calculation of the friction. For calculation purposes it was nesessary first to establish the geometry and disribution of asperities and provide its mathematical description. Two-dimensional roughness measurements were carried out by means of a contact stylus instrument in order to investigate mircogeometrical properties. Methods of mathematical statistics were used to characterize the surface roughness. The present paper follows the simplified approach of using spherical caps for describing asperity summits, and assumes a Gaussian distribution of their heights. The approximation of loaded rough surfaces was iteratively calculated until equilibrium was reached between external loading and the load-carrying capacity of asperities. The load-carrying and deformation behaviour of the surface was described by load-indentation depth curves obtained from measurements. Subsequently, also the real contact area could be determined from the approximation of both surfaces. This determination was followed by friction calculations. The deformation component of the friction force was established for two asperities, which were representative and typical for the individual surfaces (refer to Section 3.4). The adhesion component of the friction force was determined by using the shear approach (refer to Section 3.5) and the shearing-induced contact area increase ('junction growth') was considered for plastic deformations of asperities. A comparsion between the results obtained from measurements and those obtained from calculations yielded a good correspondence.

As turned out, the friction coefficient is not a constant and varies under the influence of various operating variables as a result of changed microgeometrical and material properities and of varying ambient conditions. Consequently, each operating point in the rolling contact needs to be assigned its specific friction coefficient. For further details as to friction calculations for wheel-to-rail contact reference is made to [14].

Apart from determining friction, also the wear volume and knowledge about wear processes occuring in the wheel-to-rail contact are of critical im-

portance to the development of new material and scheduling maintenance operations, such as regrinding rail sections.To this end, wear tests were performed on a roller/roller model under various predefined parameters. Experimental examinations required a new test bench be designed and built. On the new test bench the normal load was applied by means of a novel supporting roller system. This design prevented misalignment of the test rollers as turned out to be a weak point of traditional test machines of this kind. The new test bench yielded a good reproducibility of results.

The wear volume increase proved to be proportional to the increase of the acting pressure and the creep between the test rollers. Increasing the cirumferential speed of the test rollers caused a reduction in the wear volume both at the driving and at the driven roller. A significant wear reduction at the driven rail rollers could not be achieved by periodically reversing the direction of the acting friction forces; under such reversing conditions of operation a change in the wear volume at the driving rollers manufactured from wheel material could not be detected. At creep values of about 0.25 % a change between low wear and high wear levels was found for the material combination of R7 (driving) against 900A (driven).

The driving roller and the driven roller differed fundamentally in their wear processes since their subsurface material layers are exposed to different stresses and, consequently, exhibit a different level of work-hardening. Hence, wear of the driving roller, which was exposed to lower stresses and exhibits a lower surface hardness compared to the driven roller was mainly caused by tribochemical reactions, whereas fatigue and abrasion were the marked wear-reducing mechanisms at the driven roller.

Examinations as to the extent to which theoretical results can be applied to practise, revealed a good correspondence between wear-specific friction work of the model and of a real system. Additional details and comprehensive information about wear investigations performed can be found under [15].

Acknowledgements

References

1. Johnson K. L. (1985) Contact Mechanics. Cambridge University Press, Cambridge
2. Greenwood J. A., Williamson J. B. P. (1996) Contact of Nominally Flat Surfaces. Proc. of the Royal Soc. London A, 295, 300-319
3. Holland J., Deters L., Fleischer G., Bartel D., Lüneburg B., Hadler J. (1997) Gleitlagerspontanversagen. FVV-Bericht 596-2

4. Whitehouse D.J., Archard J. F. (1970) The Properities of Random Surfaces of Significance in their Contact. Proc. Roy. Soc. London A, 316, 97-121

5. Deters L., Engel S., Proksch M. (2000) Friction and Wear at Tractive Rolling Cylinders. Zamm. Zeitschrift für angewandte Mathematik und Mechanik, Berlin 80, Supplement 1, S. 37 - S. 40

6. Greenwood J. A. (1992) Contact of Rough Surfaces. In: Singer I.L., Pollock H. M.(Eds.) Fundamentals of Friction: Macroscopic and Microscopic Processes. NATO ASI Series, Series E, Applied Sciences, Vol. 220, Kluwer Academic Publishes, Dordrecht Boston London

7. Gelinck E. (1999) Mixed Lubrication of Line Contacts. PhD Thesis, University Twente, Enschede

8. Weinhauer D. (1996) Reibung und Verschleiss einer hochbeanspruchten Reibpaarung unter Mischreibungsbedingungen. Diss. Otto-von-Guericke Universität Magdeburg, Magdeburg

9. Bowden F. P., Tabor D. (1959) Reibung und Schmierung fester Körper. Springer, Berlin Göttingen Heidelberg

10. Beagley T. M. (1976) Severe Wear of Rolling/Sliding Contacts. Wear 36, 317-335

11. Zastrau B., Nackenhorst U., Jarewski J. (1999) Zur numerischen Analyse der Beanspruchung von Rad und Schiene beim Rollkontakt. Zwischenbericht zum Forschungsprojekt Za 141/5-2 in DFG - SPP "Systemdynamik und Langzeitverhalten von Fahrwerk, Gleis und Untergrund" Hamburg Dresden

12. Knothe K., Theiler A. (1998) Vergleich Radsatz-Walzenprüfstand. Sind Laborversuche auf die Realität übertragbar? Seminar Statik und Dynamik von Konstruktionen, TU Berlin, Berlin

13. Knothe K., Gueney S., Theiler A., Zhang J. (1998) Beanspruchungen von Schienenlaufflächen beim Bogenlauf. In: Tagungsband zur VDEI Fachtagung Bahn Bau '98, Berlin, October 28-30

14. Engel S.(2002) Reibungs- und Ermüdungsverhalten des Rad-Schiene-Systems mit und ohne Schmierung. Diss. Otto-von-Gurericke Universität Magdeburg, Magdeburg

15. Proksch M. (2002) Simulation des Wälzverschleissverhaltens von Rad/Schiene-Werkstoffen. Diss. Otto-von-Guericke Universität Magdeburg, Magdeburg

Model-Based Validation
within the Rail-Wheel-Subgrade Modeling

Dirk Söffker[1], Svenja Kirchenkamp[1], and Peter C. Müller[2]

[1] Gerhard-Mercator University Duisburg, Dynamics and Control,
   Lotharstr. 1, 47048 Duisburg, Germany
[2] University of Wuppertal, Safety Control Engineering,
   Gaußstr. 20, 42097 Wuppertal, Germany

Abstract. This contribution presents the idea of model-based model validation. For this purpose the PI-Observer as core of a validation scheme is introduced. Task of this observer scheme is the estimation of unknown, unmeasured interaction effects within the considered system. Therefore it is assumed, that the whole system is divided in known and unknown parts. Estimating dynamic interaction effects inner ratios within the unknown part are reconstructed. This model-based reconstruction scheme can be used for validation purposes if real measurements are available. As example the adhesion models are validated. The contact force highly depends on the kinematical contact situation and additionally on external effects. By observing the contact situation by the actual adhesion characteristic between rail and wheel can be determined.
As a first application example the simulations of the driven wheel show the strategy and the efficiency of this approach.

1 Motivation

For model validation by comparison of predicted with real values, measurements are needed. Here usually costly experimental validation especially for models with a wide application range are necessary. Further on model validation often is done using specific test-rigs to realize conditions to allow measurements, which typically can not be assigned to practical problems. Another problem is that measurements of the interesting system part and of the interaction effects to others are not available. This means that practical realization of measurements restricts the model refinements to those effects, which affects the output. The proposed method overcomes these difficulties by the application of a robust model-based technique which uses simple to realize measurements and model knowledge to predict inner values of unknown parts of the considered model.

The principle idea is illustrated by the validation of the nonlinear adhesion characteristic. The adhesion-friction micromechanism is the core of the transport mechanism of locomotion. Figure 1 gives possible contact situations of a rail-wheel contact. The area of an optimal use of the contact situation is defined due to the assumption of the maximum position of the friction-slip ratio on the left hand side of the maximum, if a maximum is available.
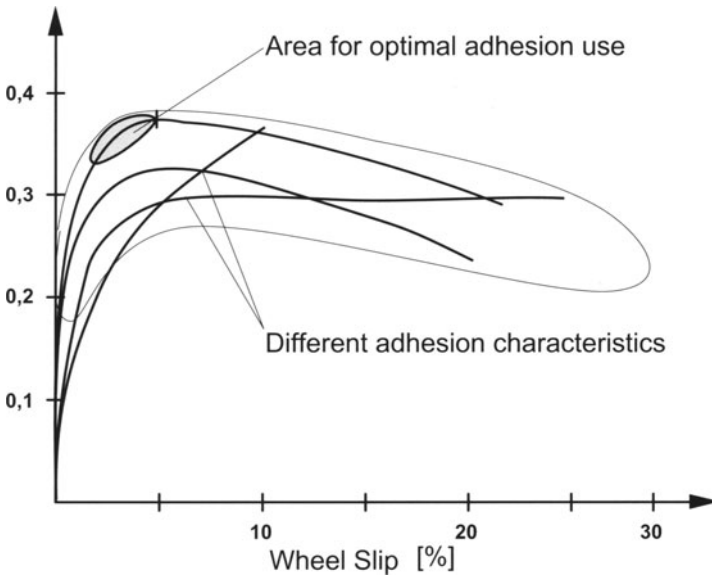
Fig. 1. Area of the optimal adhesion of the rail-wheel contact

The rail-wheel contact is highly nonlinear due to the complex geometrical contact problem and the unknown environment parameters (temperature et al.). Several papers focus to the modeling of the contact phenomena, trying to understand the effects as well [10], [21]. Furthermore the interaction between the elastic track, the elastic rail, the elastic contact itself and the wheel is of interest, because of several safety, economical and comfort aspects [18].

The problem of validation of those contact models are unavailable measurements of the contact area. The model-based model validation technique offers with the Proportional Integral (PI) Observer a possibility for validation in cases where measurements are impossible or too costly. The PI-Observer uses known model parts and easy available measurements from the known parts to estimate unknown effects within the unknown parts.

So it works as a 'virtual measurement device'. The method is illustrated in Fig. 2.

The aim of this contribution is to demonstrate the possibilities of observer-based estimation of nonlinear contact forces. Therefore it is not necessary to work with detailed models of the mechanical system itself rather than with arbitrary disturbed models to examine the robustness of the observer technique itself to its own model assumptions.
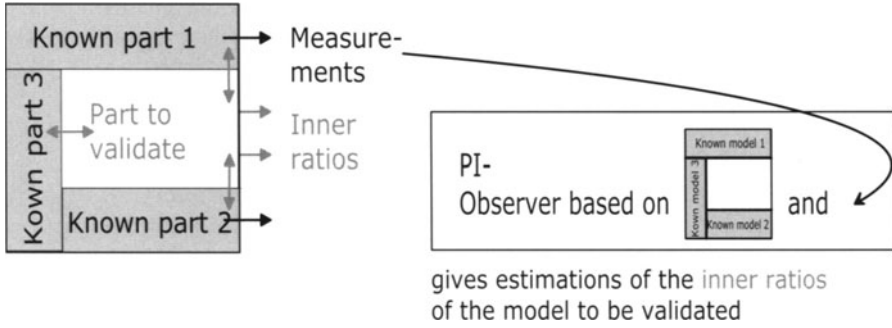
Fig. 2. The PI-Observer as a virtual measurement device

2   The Rail-Wheel Contact

Modeling of the contact between two elastic bodies is a classical problem. The solution to the normal-contact problem – the determination of the contact area and the planar pressure functions – was given by Hertz [6] . In this contribution the contact area is assumed as an ellipse, described by the radi a, b. This includes that the contact partners are shifted together with the distance

$$\delta = \left(\frac{N}{\sqrt[3]{R_a}} 2 \frac{(1-\nu)}{G} \right)^{\frac{2}{3}} \alpha_a \,, \tag{1}$$

with N acting as normal force, R_a as wheel radius, ν as the contraction number ($= 0.3$ (steal)), the shear modulus G and the coefficient α_a. The ellipse radii are connected by

$$b = a\sqrt{1 - e^2} \tag{2}$$

(cf. [11]), where e defines the eccentricity of the contact ellipse. Relations between the radii R_a, R_b of the contact partners and the eccentricity are given by

$$\frac{R_b}{R_a} = \frac{(1 - e^2)D(e)}{B(e)} \tag{3}$$

(cf. [11]), with the coefficients $D(e)$, $B(e)$ as solutions from elliptic integrals. Tabular solutions are given in [14].

In this contribution a dynamical contact is assumed, where the coefficients R_a, R_b, α_a are fixed. The coefficient α_a is choosen with $\alpha_a = 0.549$.

The contact force can be modeled assuming the theory of Kalker [9]. Here the tangential contact force is modeled using the model of Shen-Hedrick-Elkins [20]. Here only the tangential slip η is considered. The tangential contact force T_ξ is modeled by

$$T_\xi = \alpha T_\xi^{lin.} = -\alpha abGC_{11}\eta \,, \tag{4}$$

with

$$\alpha = 1 - \frac{1}{3}\left(\frac{T_\xi^{lin.}}{\mu N}\right) + \frac{1}{27}\left(\frac{T_\xi^{lin.}}{\mu N}\right)^2 \quad \text{for} \quad T_\xi^{lin.} < 3\mu N \tag{5}$$

$$\alpha = \frac{\mu N}{T_\xi^{lin}} \quad \text{for} \quad T_\xi^{lin.} \geq 3\mu N, \tag{6}$$

and C_{11} as a Kalker coefficient which is chosen fixed. The equations (1)–(6) define a piecewise constant contact force / slip relation for the elastic contact. This simplified model is sufficient to show the efficiency of this approach.

3   The Proportional-Integral-Observer

In this paper a Proportional-Integral Observer (PIO) is used, which allows the robust estimation of modeled system states, additionally the estimation of unknown inputs in desired / interesting input channels. If the nominal system behaviour can be described by a nominal system description, changes in the system structure or of system parameters can be understood as additional external inputs acting to the nominal system and representing the fault. In contrast to the Extended Kalman Filter, this procedure considers the dynamical changes both of structure and parameter. In contrast to actual works about the Unknown Input Observer (UIO) [7], this approach works in an approximated wise, but with weak conditions.

In the sequel it can be shown that the PIO allows the robust estimation of such unknown inputs interpreted as disturbances to the nominal system. The main idea of this paper is the application of this observer type to nominal known systems for fault diagnosis, in the way that the operator of a dynamical system gets a new tool looking for inner, unmeasurable states of a system. Combining the estimations of the PIO a new quality of inner informations of the faulty structure is available. The main details of PIO is already given and proved in [23]. In this paper the PIO is extended for applications as Unknown-Input Observer (UIO), applied for estimation of unknown additive inputs, like the contact forces of the rail-wheel contact.

3.1   History of Disturbance Estimation

Based on a linear and deterministic description of the plant, which describes the nominal unfaulty dynamical behaviour of the plant, the Luenberger observer can reconstruct unmeasurable states using measurements of outputs. This permits the employment of the Luenberger-observer scheme to dynamic systems of the form

$$\dot{x} = Ax + Bu, \quad y = Cx, \tag{7}$$

with the state vector x of order n, the vector of measurements y of order r_1, and the known input vector u of order m. The system matrix A, the input matrix B and the output matrix C are of appropriate dimensions. However, it is not directly applicable to nonlinear systems or systems with unknown inputs. Since the proposed type of observer beside the proportional feedback like the Luenberger-observer, also uses integral information of the estimation error, it is called PI-observer (PIO). It is known from literature, that the PI-observer design is useful for linear systems with constant disturbances [3]. Here the PI-observer is developed from another viewpoint. Continuing the ideas of Johnson [8], who introduced linear models for disturbances acting upon linear systems, and Müller [16], [17], who gave the conditions and proofs for modeling disturbances as linear models also acting upon linear systems, this paper deals with the idea of constructing a 'disturbance model' for more general use, especially to the practical case, in which no information about the disturbance, the structure of the fault resp., is available. Here the term 'disturbance model' describes the use of disturbance models describing the signal behaviour of external inputs regretted as disturbances, representing the disturbance rejection philosophy given in [8], [16] and [17]. Here this term is used only to relate the proposed development of the PI-observer to the known disturbance rejection strategy, which can be considered as a special case.

The following aspects are the points of consideration: usual Luenberger observer fails, if the system (7) is only roughly known or/and there exist additional unknown inputs caused by nonlinearities. Using known PI-observer techniques this disadvantage can be compensated, but only for piecewise constant disturbances [3]. If the unknown input is caused by modeling errors or unmodeled nonlinearities (unmodeled dynamics), this assumption is not fulfilled.

In the general case of an external input, which can not be described by a linear model extension (disturbance rejection theory), the system description (7) fails. Therefore, a more general description of such systems is given

$$\dot{x} = Ax + Bu + Nf(x, u, \mathrm{t}), \quad y = Cx \quad . \tag{8}$$

In (8) the vector function $f(x, u, \mathrm{t})$ of dimension r_2 describes in general the nonlinearities caused by the external input, unknown inputs and unmodeled dynamics of the plant and may be a nonlinear function of states, control inputs and time. The matrix N is the corresponding distribution matrix locating the unknown inputs to the system. Without loss of generality here it is assumed that the matrices N, C have full rank.

Several successful practical and theoretical applications concerning machine diagnosis [22] and also observer-based control [2], [15] are known. In all of these cases an approximation

$$f \approx Hv \tag{9}$$

of the vector of nonlinearities f (friction torques, forces caused by the crack) was used. In the theory of DRC [8], [16], [17], the linear time-invariant system

with the unknown inputs Nf caused by nonlinearities, unknown inputs or unmodeled dynamics is described by the linear exo-system

$$\dot{v} = Fv \quad . \tag{10}$$

The resulting extended linear dynamical model includes these inputs and appears as

$$\begin{bmatrix} \dot{x} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} A & NH \\ 0 & F \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u, \tag{11}$$

$$y = \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix}, \tag{12}$$

and can be used as an (extended) base building up an extended linear observer. Here the matrix N relates the fictitious approximations Hv of the unknown inputs f to the states where they appear. The signal characteristics of these inputs will be approximated by a linear dynamical system with the system matrix F. Using the extended system description (11),(12) an extended observer can be designed, so the estimate $\hat{v}$ of v represents the approximation of the disturbances, whereby $\hat{x}$ is the estimation of x.
In the applications [22], [2], [15] it is noticed that using

$$F = 0, \quad F \to 0, [\text{resp.,}] \tag{13}$$

leads to a very good reconstruction of the diagnosted nonlinearity. This means that without exact knowledge about the dynamical behaviour of the unknown inputs f ($F = 0$ represents a constant disturbance [3]), a very general approach is possible by assuming the disturbance as approximately piecewise constant (related to the 'disturbance model'-philosophy), but applying the observer scheme to applications where this assumption is not fulfilled.
By the given procedure using '$F = 0$' in the sense of disturbance model philosophy, the successfully applied scheme appears as the PI-observer and will be seen as a natural comprehensible extension of the well known Luenberger observer [23].
Figure 3 shows the structure of the observer therefore (8) is written in the following form

$$\dot{x} = Ax + Bu + b + N^*n^*(x, t). \tag{14}$$

where b is the known input which is independent from u and N^*n^* describes the unknown, external inputs.
Here in contrast to the conventional Luenberger approach a second loop with two gain matrices L_2, L_3 and integrator is used additionally.
Now, the question is how to determine the matrices L_1, L_2 and L_3 such that the corresponding PIO works well. Therefore the estimation performance is analyzed for different cases of the dynamical behaviour of the unknown input related to the nominal system.
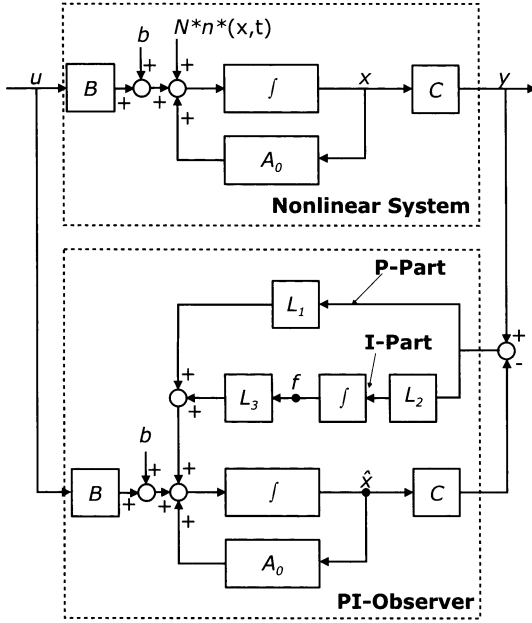
Fig. 3. Structure of the PI-Observer

3.2   Estimation Behavior

The development of the theory behind follows the works in [23], [26]. From the structure of the PI-observer depicted in Fig. 3 it follows, that the dynamics of PI-observer is described by

$$\dot{\hat{x}} = A\hat{x} + L_3\hat{f} + Bu + L_1(y - \hat{y}) \qquad\qquad \dot{\hat{f}} = L_2(y - \hat{y}) \tag{15}$$

where $\hat{y} = C\hat{x}$. Writing (8) in a matrix form gives

$$\begin{bmatrix} \dot{\hat{x}} \\ \dot{\hat{f}} \end{bmatrix} = \begin{bmatrix} A & L_3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{f} \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u + \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} (y - \hat{y}) \tag{16}$$

or

$$\begin{bmatrix} \dot{\hat{x}} \\ \dot{\hat{f}} \end{bmatrix} = \underbrace{\begin{bmatrix} A - L_1 C & L_3 \\ -L_2 C & 0 \end{bmatrix}}_{A_e} \begin{bmatrix} \hat{x} \\ \hat{f} \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u + \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} y. \tag{17}$$

Now the problem is how to design the gain matrices L_1, L_2, and L_3, such that the observer can estimate approximately the states x of the plant.

Defining the estimation error as $e(t) = \hat{x}(t) - x(t)$. Then, from (7), (8) and (17) we have that

$$\begin{bmatrix} \dot{e} \\ \dot{\hat{f}} \end{bmatrix} = A_e \begin{bmatrix} e \\ \hat{f} \end{bmatrix} \tag{18}$$

in the case of system (7), or

$$\begin{bmatrix} \dot{e} \\ \dot{\hat{f}} \end{bmatrix} = A_e \begin{bmatrix} e \\ \hat{f} \end{bmatrix} - \begin{bmatrix} N \\ 0 \end{bmatrix} f \tag{19}$$

in the case of system (8) with unknown inputs or nonlinearities. From (17) the following result can be obtained [23].

3.3 Known System without External Inputs

Theorem 1. If the pair (A, C) is observable, then there exists a PI-observer with any dynamics for the system (7), such that $\lim_{t \to \infty}[\hat{x}(t) - x(t)] = 0$ for any initial states $x(0)$, $\hat{x}(0)$ and $\hat{f}(0)$.

Proof. From the dynamics (17) of PI-observer it can be seen that the dynamics or poles of (17) can be arbitrarily assigned if and only if the matrix pair $\left(\begin{bmatrix} A & L_3 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} C & 0 \end{bmatrix} \right)$ is observable, i.e.

$$rank \left\{ \begin{bmatrix} sI - A & -L_3 \\ 0 & sI \\ C & 0 \end{bmatrix} \right\} = n + dim(\hat{f}) \tag{20}$$

holds for all $s \in \mathbf{C}$. Furthermore, the condition (19) is equivalent to

$$rank \left\{ \begin{bmatrix} A & L_3 \\ C & 0 \end{bmatrix} \right\} = n + dim(\hat{f}) \tag{21}$$

when $s = 0$ and

$$rank \left\{ \begin{bmatrix} sI - A \\ C \end{bmatrix} \right\} = n \tag{22}$$

when $s \neq 0$. The condition (21) implies that the dimension of the integrator must be less than or equal to that of the outputs. Since the matrix L_3 may be arbitrarily selected, the rank condition (22) holds if and only if

$$rank \left\{ \begin{bmatrix} A \\ C \end{bmatrix} \right\} = n \quad . \tag{23}$$

Combining the conditions (22) and (23) leads to

$$rank \left\{ \begin{bmatrix} sI - A \\ C \end{bmatrix} \right\} = n \tag{24}$$

for all $s \in \mathbf{C}$, i.e. (A, C) is observable.

A main motivation to study the PI-observer is to reconstruct the states of the system (8) with nonlinearities. The following two theorems give the results in case of the system (8).

3.4   Known Systems with Constant External Inputs

Theorem 2. Assume that $\lim_{t \to \infty} f(x, u, t)$ exists. Then, there exists a PI-observer with any dynamics for the system (8), such that $\lim_{t \to \infty} [\hat{x}(t) - x(t)] = 0$ for any initial states $x(0)$, $\hat{x}(0)$ and $\hat{f}(0)$ if (A, C) is observable and

$$rank \left\{ \begin{bmatrix} A & N \\ C & 0 \end{bmatrix} \right\} = \mathrm{n} + \mathrm{r}_1 \tag{25}$$

Proof. Using the construction method, we prove Theorem 2. Let $L_3 = N$. Then, the dynamics (19) of the estimation error of PI-observer (17) becomes

$$\begin{bmatrix} \dot{e} \\ \dot{\hat{f}} \end{bmatrix} = A_e \left\{ \begin{bmatrix} e \\ \hat{f} \end{bmatrix} - \begin{bmatrix} 0 \\ I \end{bmatrix} f \right\} \tag{26}$$

where $A_e = \begin{bmatrix} A - L_1 C & N \\ -L_2 C & 0 \end{bmatrix}$. Similarly with the proof of Theorem 1, the eigenvalues of the matrix A_e can be arbitrarily assigned by the matrices L_1 and L_2 if and only if the matrix pair $\left(\begin{bmatrix} A & N \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} C & 0 \end{bmatrix} \right)$ is observable, i.e.

$$rank \left\{ \begin{bmatrix} sI - A & -N \\ 0 & sI \\ C & 0 \end{bmatrix} \right\} = \mathrm{n} + \mathrm{r}_1 \tag{27}$$

holds for all $s \in \mathbf{C}$. This condition is equivalent to

$$rank \left\{ \begin{bmatrix} A & N \\ C & 0 \end{bmatrix} \right\} = \mathrm{n} + \mathrm{r}_1 \tag{28}$$

when $s = 0$ and

$$rank \left\{ \begin{bmatrix} sI - A \\ C \end{bmatrix} \right\} = \mathrm{n} \tag{29}$$

when $s \neq 0$. This implies that under the conditions in theorem 2 the dynamics of PI-observer (16) for the system (8) can be arbitrarily assigned. Therefore, the eigenvalues of A_e can be arbitrarily placed at any locations in the left-half complex plane when the conditions in theorem 2 are satisfied. This means that the dynamics (26) is stabilizable by means of the matrices L_1 and L_2. When the dynamics (26) is asymptotically stable, its solution will converge to the equilibrium. Then, from (26) it can be easily seen that

$$\lim_{t \to \infty} \begin{bmatrix} e(t) \\ \hat{f}(t) \end{bmatrix} = \begin{bmatrix} 0 \\ \lim_{t \to \infty} f(x, u, t) \end{bmatrix} \quad . \tag{30}$$

3.5   Known Systems with Arbitrary External Inputs

Theorem 3. Assume that $f(x, u, \mathrm{t})$ is bounded. Then, there exists a high-gain PI-observer for the system (8) such that $\hat{x}(\mathrm{t}) - x(\mathrm{t}) \to 0$ ($\mathrm{t} > 0$) for any initial states $x(0)$, $\hat{x}(0)$ and $\hat{f}(0)$ if
     1) (A, C) is observable, which includes

$$rank \left\{ \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\mathrm{k}-1} \end{bmatrix} \right\} = \mathrm{n} \quad, \tag{31}$$

where k is the observability index of (A, C),
     2) $rank \left\{ \begin{bmatrix} A & N \\ C & 0 \end{bmatrix} \right\} = \mathrm{n} + \mathrm{r}_1$; and
     3) $CN = 0$

Proof. Let $L_3 = N$. Then, analogously with the proof of theorem 2, it is easily verified that the dynamics of PI-observer (17) for the system (8) can be arbitrarily assigned by means of the matrices L_1 and L_2 if the conditions 1) and 2) in theorem 3 are satisfied.
Under the selection of L_3 the dynamics (19) of the estimation error becomes (26). When A_e is stable, the solution to (26) will be also bounded if $f(x, u, t)$ is bounded. Let $\begin{bmatrix} L_1 \\ L_2 \end{bmatrix} = \rho_1 \begin{bmatrix} \tilde{L}_1 \\ \tilde{L}_2 \end{bmatrix}$. Then, (26) may be written as

$$\frac{1}{\rho_1} \begin{bmatrix} \dot{e} \\ \dot{\hat{f}} \end{bmatrix} = \frac{1}{\rho_1} \begin{bmatrix} A & N \\ 0 & 0 \end{bmatrix} \begin{bmatrix} e \\ \hat{f} \end{bmatrix} - \begin{bmatrix} \tilde{L}_1 \\ \tilde{L}_2 \end{bmatrix} Ce - \frac{1}{\rho_1} \begin{bmatrix} N \\ 0 \end{bmatrix} f \tag{32}$$

From (32) it follows that

$$Ce = 0 \tag{33}$$

for $\rho \to \infty$. Differentiating (33) and using (19) give

$$C\dot{e} = C(A - L_1 C)e + CN(\hat{f} - f) \tag{34}$$

From the condition 3) and (33) we have

$$CAe = 0 \quad . \tag{35}$$

In the same way under the condition 1) we can obtain

$$CA^i e = 0 \quad i = 0, 1, ..., \mathrm{k} - 1 \tag{36}$$

Then from (33), (35) and (36) it follows that

$$e = 0 \tag{37}$$

due to condition 1). Substituting (37) into (19) gives

$$\hat{f} - f = 0 \tag{38}$$

because of the full-column rank of N. Equations (37) and (38) mean that the estimates $\hat{x}$ and $\hat{f}$ of the PI-observer (19) converge to the states x and the unknown inputs f of the system (8) when ρ_1 goes to the infinity. This shows that $\hat{x}$ and $\hat{f}$ may approximate x and f in the case of high gains.

In [23], [25] furthermore it is shown, that this type of observer also can be applied in general to systems not completely known with unknown additive inputs. In contrast to the mentioned works in the meantime the condition for the application of the PI-Observer to such structures is corrected to the conditions

$$\rho_1 \to \infty \quad \text{and} \tag{39}$$

$$\frac{\rho_2}{\rho_1} \to \infty \quad \text{with} \quad L_3 = \rho_2 N. \tag{40}$$

which gives theoretical hints to understand the observed success of the observer technique in robotics and machine-dynamics. In this application the PIO is applied to known systems with arbitrary external inputs.

4   Modeling and Simulation

Figure 4 illustrates the system to be considered: a torsion-stiff wheelset with linear springs c_{wh}, c_{wv} and dampers d_{wh}, d_{wv} for horizontal and vertical degrees of freedom. The electric drive is coupled with an elastic torsion spring-damper combination c_{MW}, d_{MW} to the wheelset. The other constants are the motor inertia Θ_M, the rotational inertia of the wheelset and drive Θ_W, and the related mass of the wheelset m_W. The modeled degrees of freedom are the motor angle φ_M, the wheel angle φ_W, the horizontal displacement of the wheelset u_{wx} and the vertical displacement of the wheelset u_{wz}. Modeled, but not given in the illustration fig. 4 are also the modal displacements of the rail u_{g1}, u_{g2} and the angles for the spatial orientation of the wheelset φ_1, φ_2.

The equations of motion are given by

$$\ddot{\varphi}_M + \frac{d_{MW}}{\Theta_M}(\dot{\varphi}_M - \dot{\varphi}_W) + \frac{c_{MW}}{\Theta_M}(\varphi_M - \varphi_W) = M(\mathrm{t}) \tag{41}$$

$$\ddot{\varphi}_W + \frac{d_{MW}}{\Theta_M}(\dot{\varphi}_W - \dot{\varphi}_M) + \frac{c_{MW}}{\Theta_M}(\varphi_W - \varphi_M) = \frac{2rT_\xi}{\Theta_W} \tag{42}$$

$$\ddot{u}_{wx} + \frac{2d_w}{m_w}\dot{u}_{wx} + \frac{2c_w}{m_w}u_{wx} = \frac{2T_\xi}{m_w} \tag{43}$$

$$\ddot{u}_{wz} + \frac{2d_w}{m_w}\dot{u}_{wz} + \frac{2c_w}{m_w}u_{wz} = \frac{2N}{m_w} - g \tag{44}$$
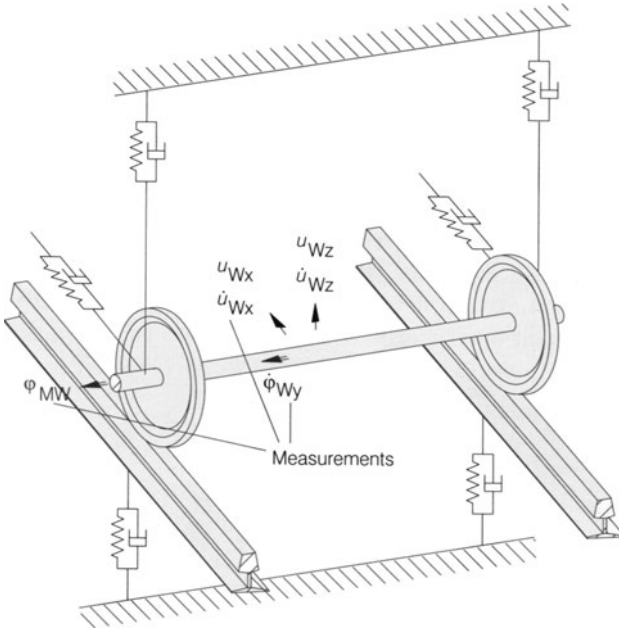
Fig. 4. Illustration of the wheelset to be considered

with the time-dependent normal contact load N, $\dot{N}$. The equations are coupled with the equations for the (modal) rail vibration and mainly due to the kinematic equations for the tangential contact force. The corresponding equation for the slip η is in absence of the absolute velocity of the bogie in the shown simulation example

$$\eta = 1 - \frac{\dot{\varphi}_W}{\dot{\varphi}_M} - \frac{\dot{u}_{wx}}{r\dot{\varphi}_M} \, . \tag{45}$$

This definition denotes some kind of averaging slip, and has to used without loss of generality to get not only the elastic slip considering the elastic horizontal displacement of the wheelset. For practical purposes (with the possibility to measure the absolute speed) the usual definitions using the absolute speed v_0 can be used.

It should be noted, that there also exists a kinematical coupling between the rail, the elastic contact and the disturbance height Δz

$$d = u_{g1} + u_{g2} - u_{wz} + \Delta z \tag{46}$$

of the rail. For the following simulations Δz (as a stochastic value) works as an exitation. The whole model results as a nonlinear system with 10 elastic degrees of freedom. The main features of this model are taken from [13].

To get realistic adhesion characteristics the friction coefficient μ is stochasti-

cally modified. So a much more realistic characteristic as given by the Shen-Hedrick-Elkins model (4)–(6) as shown in Fig. 5 is available.
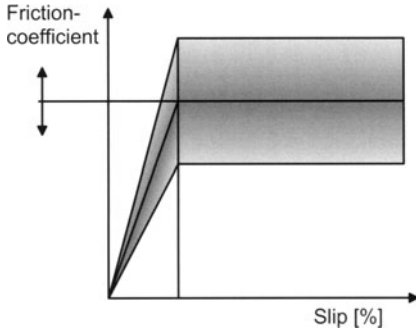
Fig. 5. Adhesion characteristic with stochastically modified friction coefficient

It should be noted that the PIO-design is independent from this structural and parametrical information. For the PIO only the known parts of (41)–(43) are used, the parts including T_ξ and the other equations are only used for the system simulation model.

4.1 Simulation Results

For the simulations the input is the torque $M(\mathrm{t})$ as a step function.
In Fig.1 the results for $M(\mathrm{t}) = 5 \cdot 10^4$ Nm and for the PI-Observer using the two measurements φ_M and u_{wx} are depicted. The upper plot shows the time behavior of the tangential contact force T_ξ (marked as o) and its estimation (marked as +). Then the adhesion characteristic – that is T_ξ/N depending on the slip η – is shown. Here the absolute value of η is used to get the rolling condition $\eta = 1$ with the used slip definition (45). This first plot shows that the observer works very well in the case of using two measurements.

Next the observer is used with only the measurement of φ_M and the same torque size as in the first simulation. Fig. 4 shows the results of that simulation. Again the estimation of the time behavior of the contact force works quite well but the adhesion characteristic is only roughly estimated.

For the same settings as in the first simulation but with a torque twice as there the results are given in Fig. 5. This shows that the success is independent of the height of the input.

The applied torque for the following simulations is again $M(\mathrm{t}) = 5 \cdot 10^4$ Nm. For an observer with two measurements with additional noise the results of
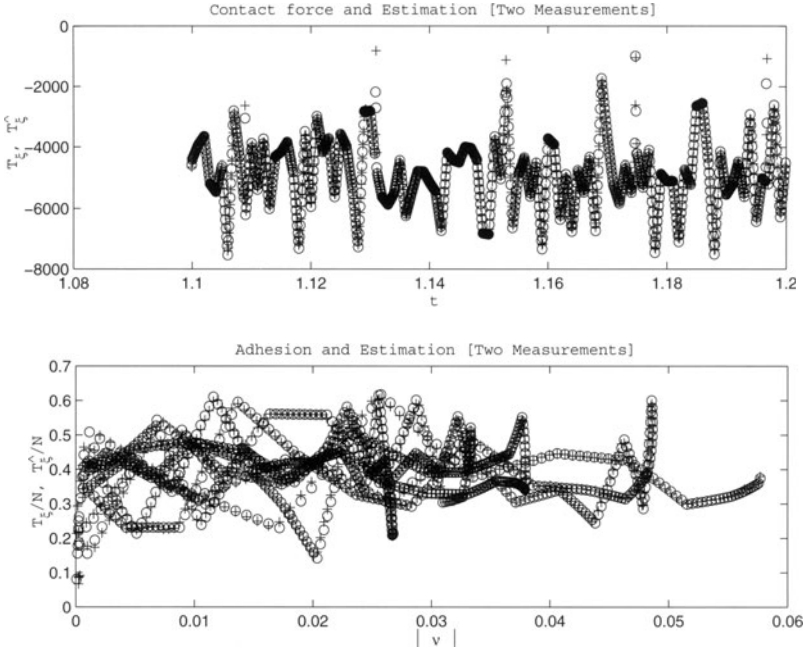
Fig. 6. a) Time behavior of the contact force and b) adhesion characteristic
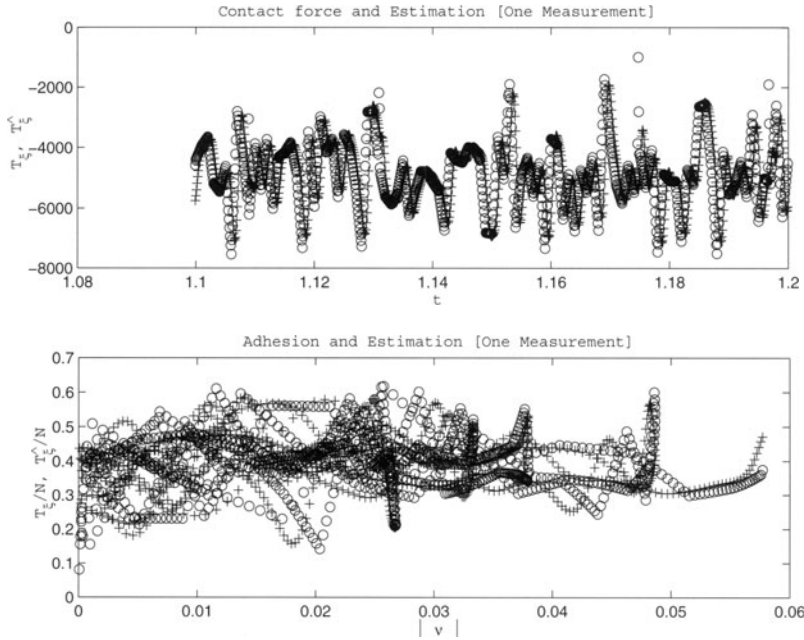
Fig. 7. a) Time behavior of the contact force and b) adhesion characteristic
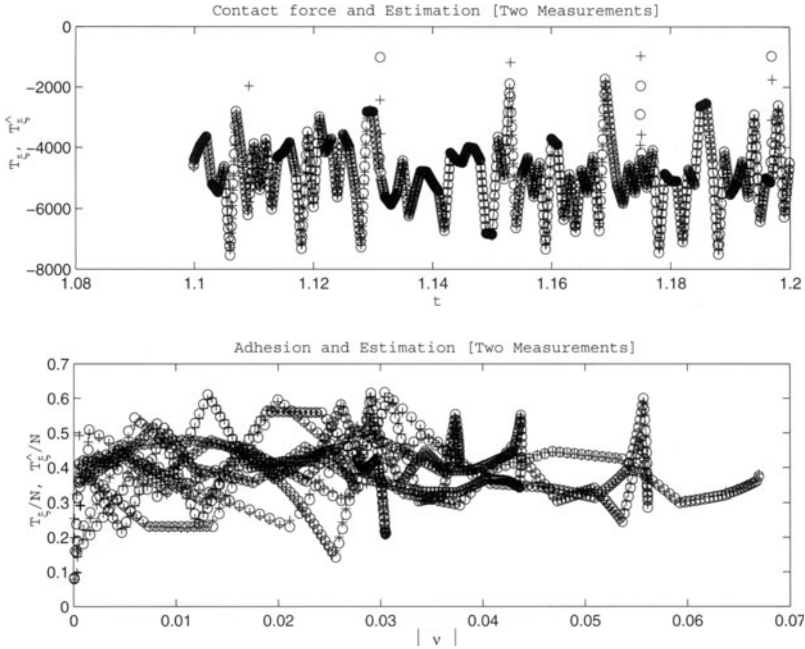
Fig. 8. a) Time behavior of the contact force and b) adhesion characteristic

the simulation are shown in Fig. 2. The estimation of the time behavior as well as the estimation of the adhesion characteristic is close to the real run of the curve.

Figure 3 shows the result for an observer with two measurements with higher noise than shown in the results given with Fig. 2. An additional difference to the case before is the consideration of a shorter time slice of the simulation. This resolution displays less good estimations of the curves.

5   Concluding Remarks and Future Aspects

In this contribution the application of the Proportional-Integral Observer (PIO) for model-based model validation purposes within the rail-wheel-sub-grade modeling is shown. As a first example the reconstruction of the contact behavior (usually modeled by contact models) is realized and therefore a scheme for contact model validation is implemented. Using a theoretical model of a elastic supported wheelset with a nonlinear contact model it can be shown by simulations that the linear PIO-scheme is able to estimate the nonlinear tangential contact forces. This inside view into the unmeasurable contact situation gives the base for advanced adhesion control strategies.
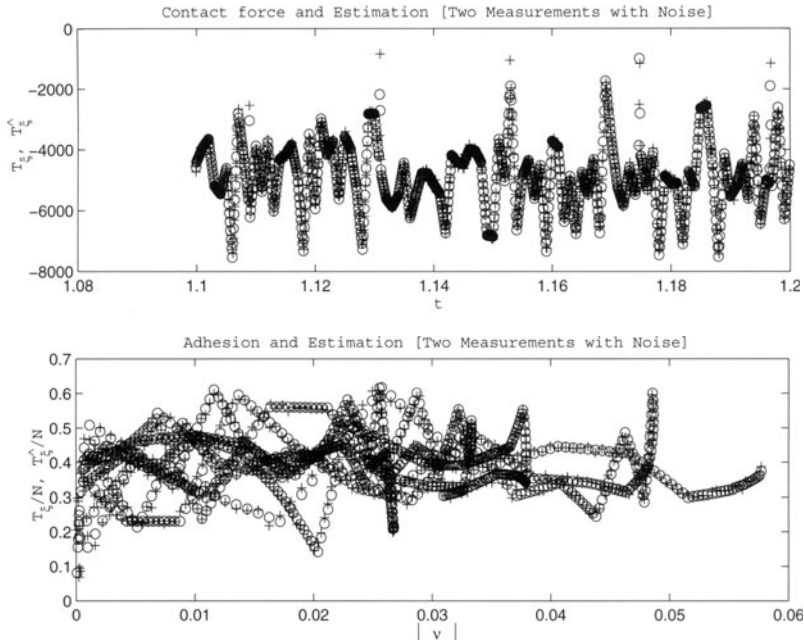
Fig. 9. a) Time behavior of the contact force and b) adhesion characteristic
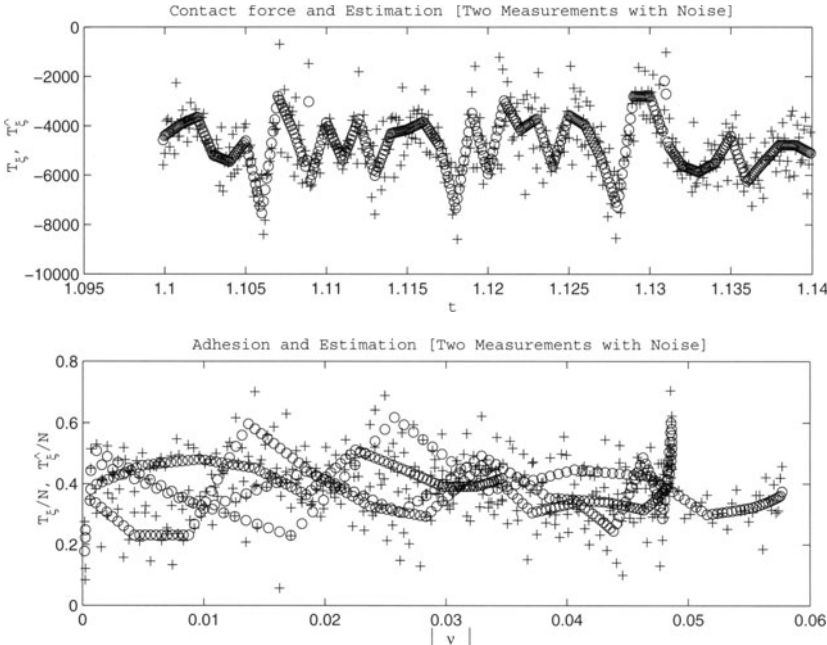
Fig. 10. a) Time behavior of the contact force and b) adhesion characteristic

The next step is to replace the simulation model by a real system and to compare the observer values originated with real measurements with the model values of the model to be validated. In other applications this comparison shows satisfying result [1].

Further as an other example the detection of voids within the subsoil is aimed. In this additional application the observer estimates the stiffness coefficients within the subsoil with the aid of measurements of the normal forces and the sleeper angle.

Acknowledgement

References

1. Abicht, C.; Bormann, J.; Müller, P.C.; Söffker, D.; Ulbrich, H. (2001) Model-Based Estimation of Impact Forces affecting Elastic Structures: Simulation and Experiment. Proc. 18th ASME-DECT Biennial Conference on Mechanical Vibration and Noise, Symposium on Vibration Including Friction and Impacts, Pittsburgh, Pennsylvania, USA, September 9-13, 2001, 6 pages
2. Ackermann, J. (1989) Positionsregelung reibungsbehafteter elastischer Industrieroboter. VDI-Fortschrittberichte, Reihe 8, Nr. 180, VDI-Verlag, Düsseldorf
3. Anderson, B.D.O.; Moore, J.B. (1989) Optimal Control - Linear Quadratic Methods. Prentice Hall, New York
4. Buscher, M.; Pfeiffer, R.; Schwartz, H.J. (1993) Radschlupfregelung für Drehstromlokomotiven. eb - Elektrische Bahnen 91, 1993, Seiten 163-178
5. Hahn, K.; Hase, K.R.; Sommer, H. (1993) Fortschritte bei der Kraftschlußausnutzung für die Hochgeschwindigkeits- und Schwerlasttraktion. ETR 42, 1993, Seiten 67-74
6. Hertz, H. (1882) Über die Berührung fester, elastischer Körper und über die Härte. Verhandlungen des Vereins zur Förderung des Gewerbefleißes, Leipzig, Nov. 1882
7. Hou, M.; Müller, P.C. (1992) Design of Observers for Linear Systems with Unknown Inputs. IEEE Trans. Aut. Control, Vol. 37, 1992, pp. 871-875
8. Johnson, C.D. (1976) Theory of Disturbance-Accomodating Controllers. In: Leondes, C.T. (Ed.): Control and Dynamic Systems. Academic Press, Vol. 12, 1976, pp. 387-489
9. Kalker, J.J. (1990) Three-Dimensional Elastic Bodies in Rolling Contact. Kluwer Academic Publishers. Dordrecht, Boston, London
10. Kalker, J.J.; Dekking, F.M.; Vollebregt, E.A.H. (1997) Simulation of Rough, Elastic Contacts. Journal of Applied Mechanics, Vol. 64 (1997), pp. 361-368
11. Knothe, K. (1992) Lateraldynamik von Schienenfahrzeugen. Vorlesungsskript TU Berlin, Wintersemester 1992/93

12. Lang, W.; Roth, G. (1996) Optimale Kraftschlußausnutzung bei Hochleistungs-Schienenfahrzeugen. ETR 42, 1996, Seiten 61-66

13. Lange, M.; Groß-Thebing, A.; Stiebler, M.; Knothe, K. (1995) Simulations-modell für Vollbahn-Antriebssysteme zur Untersuchung der ausnutzbaren Zugkraft. Arbeitsbericht zum DFG-Forschungsvorhaben Kn/St 132/20, 1993 bis 1995

14. Lurje, A.I. (1963) Räumliche Probleme der Elastizitätstheorie. Akademie Verlag, Berlin

15. Neumann, R.; Moritz, W., (1990) Observer-based joint controller design for a robot for space operation. In: Proc. Eight CISM-IFTOMM Symp. Theory of Robots and Manipulators, Ro.Man.Sy, Krakau, Poland, July 1990

16. Müller, P.C.; Lückel, J. (1977) Zur Theorie der Störgrößenaufschaltung in linearen Mehrgrößenregelsystemen. Regelungstechnik, 25, 1977, Seiten 54-59

17. Müller, P.C. (1989) Indirect Measurement of Nonlinear Effects by State Observers. IUTAM Symposium on Nonlinear Dynamics in Engineering Systems, University of Stuttgart, FRG, Aug. 21-25, 1989

18. Ripke, B. (1995) Hochfrequente Gleismodellierung und Simulation der Fahrzeug-Gleis-Dynamik unter Verwendung einer nichtlinearen Kontakt-mechanik.
    VDI-Fortschrittberichte, Reihe 12, Nr. 249, Düsseldorf

19. Schreiber, R.; Kögel, R. (1996) Identifikationsmethode zur Bestimmung der Adhäsion zwischen Rad und Schiene. ZEV+DET Glas. Ann. 120, 1996, Seiten 48-54

20. Shen, Z.Y.; Hedrick, J.K.; Elkins, J.A. (1983) A comparison of alternative creep force models for rail vehicle dynamic analysis. Proc. of 8th IASVD Symposium, MIT, Cambridge, August 15–19, 1983

21. Stiebler, M.; Knothe, K.; Dreimann, K. (Hrsg.) (1996) Tagungsband des Workshops 'Simulation und Praxis der Kraftschlußausnutzung von Hochleistungs - Triebfahrzeugen. Berlin, 27./28. Juni 1996

22. Söffker, D.; Bajkowski, J.; Müller, P.C. (1993) Crack Detection in Turbo Rotors - A New Observer - Based Method. ASME Journal of Dynamic Systems, Measurements, and Control, 3, 1993, pp. 518 - 524

23. Söffker, D.; Yu, T.J., Müller, P.C. (1995) State Estimation of Dynamical Systems with Nonlinearities by using Proportional - Integral Observer. International Journal of System Science. Vol. 26, 1995, No. 9, pp. 1571-1582

24. Söffker, D. (1995b) A New Model - Based Tool for Failure Detection and Isolation in Machine- and Rotordynamics. Proc. ASME 12th RSAFP Conf., Sept. 1995, Boston, USA

25. Söffker, D. (1996) Zur Modellbildung und Regelung längenvariabler, elastischer Roboterarme. Dissertation Universität-GH Wuppertal 1995, auch als: VDI-Fortschrittberichte, Reihe 8, Nr. 584, 1996

26. Söffker, D. (1999) Observer-based measurement of contact forces of the nonlinear rail-wheel contact as a base for advanced traction control. In: Wallaschek, J., Lückel, J., Littmann, W. (Eds.) Mechatronics and Advanced Motion Control, HNI-Verlagsschriftenreihe, Band 49, pp. 305–320

27. Zhang, J.; Knothe, K. (1995) Berechnungsmodell für Tangentialschlupfkräfte beim Kontakt in der Hohlkehle. ILR-Mitteilung 294 (1995), Institut für Luft- und Raumfahrt, TU Berlin, 1995

Part IV:

Track Dynamics

Monitoring the Dynamics of Railway Tracks by Means of the Karhunen–Loève–Transformation

Ute Feldmann, Edwin Kreuzer, Ferdinand Pinto, and Volker Schlegel

Mechanics and Ocean Engineering,
Technical University of Hamburg–Harburg,
Eißendorfer Straße 42, D-21071 Hamburg, Germany

Abstract. Maintenance of the quality of the railway track is crucial for the safety and the comfort of travelers on high speed trains. We intend to contribute a new method to detect deterioration of the track by means of a constant monitoring of the dynamic vehicle response to track features. In order to separate significant information from the enormous amount of data collected from measurements the Karhunen-Loève-Transformation (KLT) is employed. It is shown that KLT is an efficient tool for information compression and for displaying the severity of a track failure. Our project was worked out in close cooperation with Deutsche Bahn AG (German Rail).

1   Introduction

High speed traffic on the Deutsche Bahn AG (German Rail) network has continuously increased in recent years. This additional use of the network raises questions about the maintenance of the quality of the trackage, which is essential for the safety and comfort of high speed travelers.

It is known that the more frequently applied dynamic loads due to the interaction between railway vehicles, the track and the sub-grade can lead to degradation of track quality and, sometimes, premature track failure. To prevent such failures or degradation in ride quality, the Deutsche Bahn AG performs expensive, frequent inspections of the railway tracks by means of special trains designed to perform measurements on the trackage.

Track geometry is assessed by bi-monthly measurements with an especially designed car filled with measurement systems at 200 km/h. In addition, vehicle reactions are measured with the ICE-S (Intercity Express-S), at high speed, every four months, Fig. 1.

Decisions about what maintenance operations are necessary, and when, are based on empirically established threshold values of some parameters of track geometry. So far, however, there are no reliable means to predict actual rail failure based on variations of track geometry or the actions of excessive forces on the track by the train. Thus, it is difficult to schedule maintenance because one cannot predict where a failure will occur. Furthermore, there is no continuous assessment of small changes of track features, in an attempt to predict when the relevant parameters will reach limiting values.
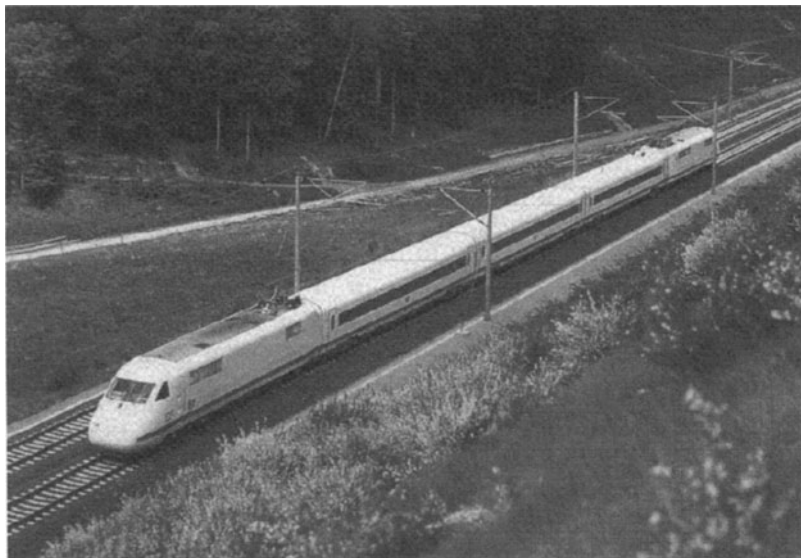
Fig. 1. Special measurement train ICE-V, predecessor of ICE-S

We propose a program to monitor track condition constantly, at low cost, on regularly scheduled trains. It is our belief that one can correlate the dynamic response of a vehicle with changes in track and support geometry, and thus we can assess track conditions by regularly measuring vehicle behavior.

We start with a description of the problem from a dynamic system point of view and present first results of measurements from a real railway track. Then we turn to a signal processing point of view. We present the Karhunen–Loève–transformation as a tool for information compression for the expected huge amount of data resulting from the constant monitoring. Finally, we demonstrate the successful application of the Karhunen–Loève–transformation to data obtained in a small scale laboratory experiment.

2   Dynamic System

The mechanical model we have used for the theoretical investigations is shown schematically in Fig. 2. The bogie of an ICE train consists of two wheel-sets. The distance between these wheel-sets is, however, small compared with the length of the coach and, therefore, we merged them to a single wheel-set. As has been described in [2], with such a model we can expect fairly good correlation between measured loads on a train and simulations based on this model.

Although this model neglects the contact dynamics between the wheels and the rails, and reduces the whole track structure (rail, rail pad, sleeper, ballast, sub-ballast, sub-grade) to a single spring damper system, the simulation
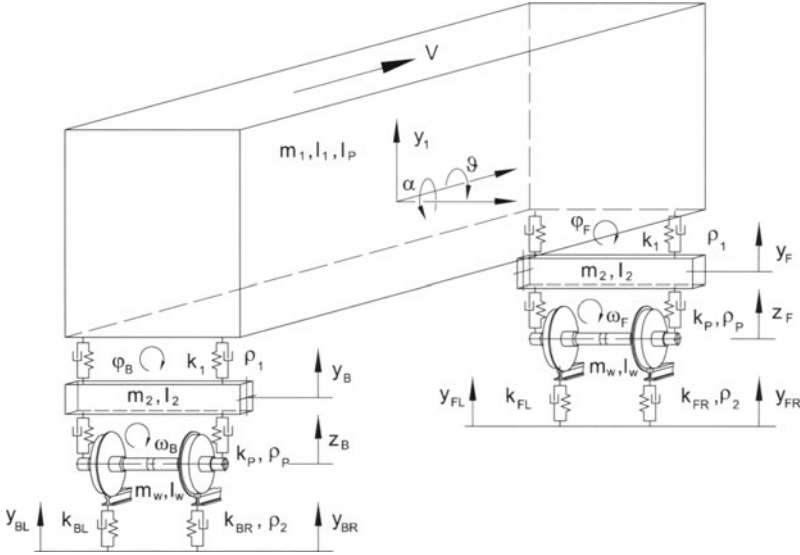
Fig. 2. Multibody system model: coach, bogie, wheel-set, primary and secondary suspension and track

based on this model allowed a fairly good prediction of track settlement due to accumulated dynamic loads. Therefore, we choose this model for our considerations.

In [2] it was found that spatial stiffness variations along the track play an important role for differential track settlement (which results in a rough vertical profile or geometry). Thus, it would be highly desirable to reconstruct the track stiffness from the dynamic response of the vehicle along the track.

Let us consider the vertical profile and the varying stiffness along the track as input u to the dynamic system Σ and the measured signals within the vehicle as output y. The task described above requires an inversion of the dynamic system, Fig. 3. The output of the inverse system u' asymptotically copies the original input signal u if and only if the inverse system has unique asymptotic behavior [1].
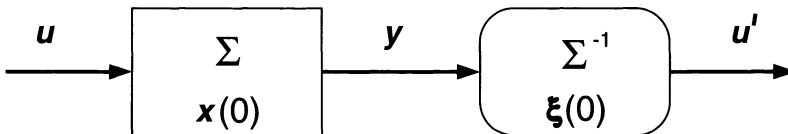
Fig. 3. Inversion of a dynamic system: the output y of the system becomes the input of the inverse system. $x(0), \xi(0)$ are the initial states of the system and its inverse

For nonlinear systems, inversion is solved in case they are *input-linear* [1], i.e., that the input u appears only as a linear term in the equations of motion. Summing up all basic equations results in $\dot{x} = f(x) + g(x) \cdot u$. Our system is, however, not input-linear, as products of the varying stiffnesses and the vertical profile occur in the equations of motion. Summing up all basic equations results in

$$
\begin{bmatrix}
m_1\ddot{y}_1 \\
I_1\ddot{\vartheta} \\
I_p\ddot{\alpha} \\
m_2\ddot{y}_F \\
I_2\ddot{\varphi}_F \\
m_2\ddot{y}_B \\
I_2\ddot{\varphi}_B \\
m_w\ddot{z}_F \\
I_w\ddot{\omega}_F \\
m_w\ddot{z}_B \\
I_w\ddot{\omega}_B
\end{bmatrix}
= f(x,\dot{x}) +
\begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
k_{FL}(t)y_{FL}(t) + k_{FR}(t)y_{FR}(t) \\
k_{FL}(t)ly_{FL}(t) - k_{FR}(t)ly_{FR}(t) \\
k_{BL}(t)y_{BL}(t) + k_{BR}(t)y_{BR}(t) \\
lk_{BL}(t)y_{BL}(t) - lk_{BR}(t)y_{BR}(t)
\end{bmatrix}
\tag{1}
$$

with the vector of generalized coordinates

$$
x = \begin{bmatrix} y_1 & \vartheta & \alpha & y_F & \varphi_F & y_B & \varphi_B & z_F & \omega_F & z_B & \omega_B \end{bmatrix}^T .
$$

The generalized coordinates together with parameters of the model are shown in Fig. 2, in addition, l is the distance between the rails. The input of the vehicle multibody system can be viewed as an excitation by means of a track which moves underneath the non moving vehicle. This results in time dependent functions of the track stiffness k_{ij} and the vertical profile y_{ij}, $i = F, B$, $j = R, L$. They are related to their spatial variation via the velocity of the vehicle. We do not present the vector $f(x,\dot{x})$ explicitly which summarizes the forces and moments of the multibody system imposed by the springs and dampers which are either linear or nonlinear.
It is an open question whether such *input-quadratic* systems are dynamically invertible.
Since we cannot infer from system inversion the track features (stiffness and vertical profile) explicitly, we want to detect the variation or deterioration of the track implicitly via the variation of the vehicle response. That means that we expect the vehicle responses at subsequent journeys under the same conditions (track features) to be the same. Thus, we expect variations of the vehicle response to result from actual variations of track features.

This is not necessarily the case because in a chaotic system, e.g., departing from different initial states would produce diverging responses to the same input signal. Thus, we require that the dynamic system have *unique asymptotic behavior* (UAB), i.e., the system response to the *same* input signal, but possibly departing from different initial states, converge to each other.

The UAB can be checked by considering the difference between any two solutions. Thus, one must prove the asymptotic stability of the origin of the difference system,

$$\Delta \ddot{x} = \ddot{x}_1 - \ddot{x}_2 = f(x_1, \dot{x}_1) - f(x_2, \dot{x}_2). \tag{3}$$

This nonlinear difference system can often be put into a linear time-varying form. The asymptotic stability of this system can be proven by the application of the Kalman–Yacubovitch Lemma [5]. The proof of UAB of our system is difficult because of the high dimension of the state space. It is the subject of ongoing research.

Next we present some measurement results in order to check UAB with real world data.

3 Measurements

We collected bi-weekly measurements with a specific vehicle on a selected piece of track in order to evaluate continuous monitoring. We installed an inertial measurement unit (IMU) from IMAR Company, Germany, on the floor of the coach. The IMU measures three translational accelerations and three angular velocities. In this way all six degrees of freedom are determined. As we are especially interested in the effects on the vertical dynamics of the track, we concentrate on measurements of the pitch angle α and the vertical displacement y_1. See Fig. 4 for a representative sample of the pitch angular velocity $\dot{\alpha}$ of two subsequent journeys along the same track at the same speed (200 km/h) and under presumably the same track conditions. The signal–to–distance correspondence along the track was determined by means of an additionally recorded signal (50 pulses/m) and the reproducibly unique yaw–signal. As the two signals in Fig. 4 are hardly distinguishable, the dynamic system actually seems to have the UAB property.

Figure 5 shows the measured pitch angular velocity $\dot{\alpha}$ and vertical acceleration $\ddot{y}_1$ of the coach in higher resolution with respect to the distance. Again the measurements of both travels nearly coincide. The signals represent the response of the coach to the onset of a curve at km 21.2 of the selected piece of the track. Apparently, an impact excites some oscillations, which are clearly visible. They have a frequency of about 1 Hz (speed was 160 km/h), which presumably corresponds to an eigenfrequency of the system (in case it is linear).

It is our belief that changes of the track, e.g., changes of the magnitude of the impact, will cause noticeable changes of the response of the coach.
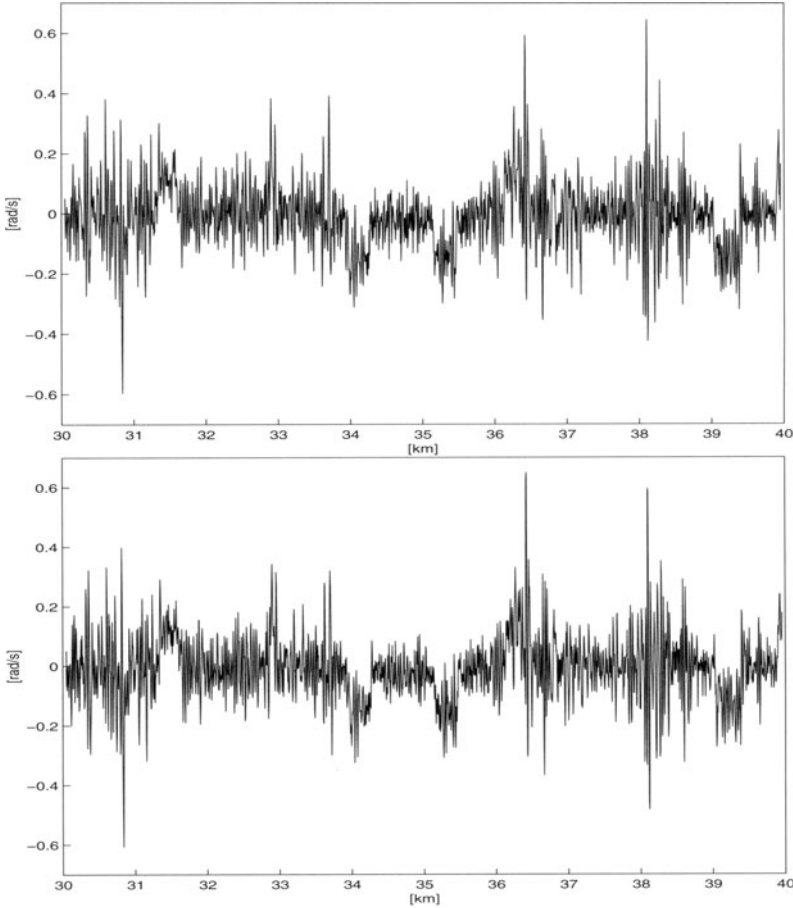
Fig. 4. Two measurements of pitch angular velocity $\dot{\alpha}$ along the same track.

Our bi-weekly measurements cover so far only a short period of time and we do not have enough data to estimate whether small deteriorations of the track are actually detectable in the coach response. Nevertheless, we performed small scale laboratory experiments, whereby it is easy to collect as much data as desired and to simulate deteriorations of the track by appropriate manipulations.

It is clear that, however, constant monitoring will result in a collection of a substantial amount of data, and will require an efficient compression and evaluation process. We propose to apply the *Karhunen–Loève–Transformation* (KLT) for this purpose. For the analysis of spatio-temporal dynamics KLT, sometimes called proper orthogoal decomposition, has been recently used as an efficient tool for data processing in several disciplines. We briefly present the basic idea of the KLT as well as its essential ingredients and then demonstrate its successful application to data of our laboratory experiments.
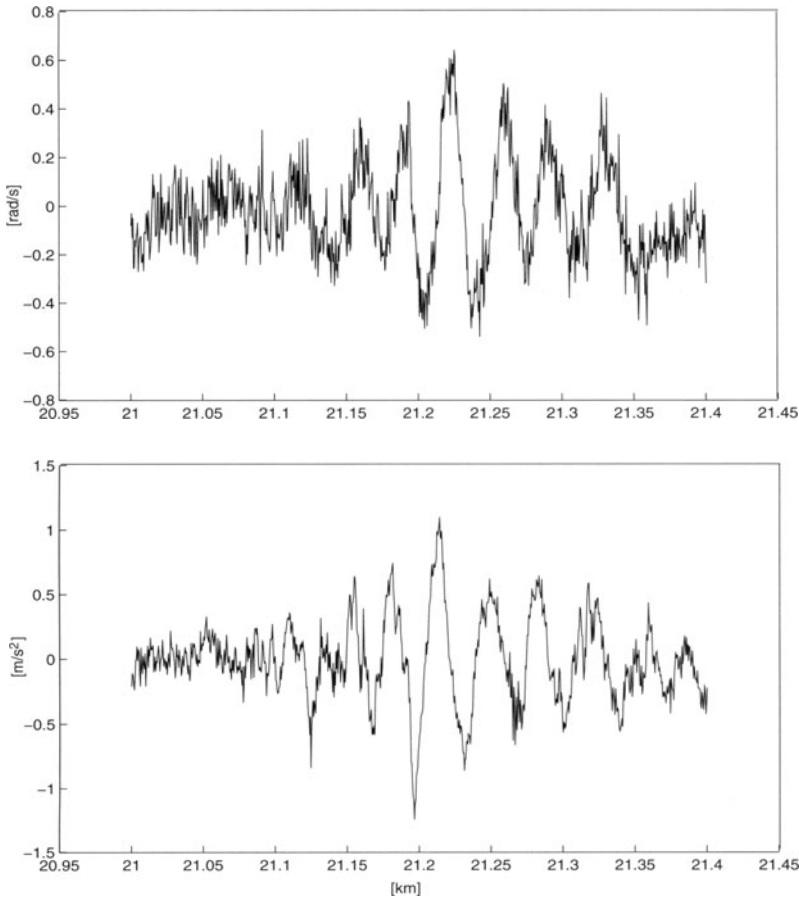
Fig. 5. Above pitch angular velocity $\dot{\alpha}$ and below vertical acceleration $\ddot{y}_1$ of the coach response to the onset of a curve (at km 21.2)

4   Karhunen–Loève–Transformation

The Karhunen–Loève–Transformation (KLT) or Proper Orthogonal Decomposition originally stems from the field of stochastic processes. A well written comprehensive presentation can be found in [3] and an application to a mechanical system is discussed in [4]. The KLT detects basic components of a set of stochastic signals by means of its covariance function. An aim can be to separate significant components from less significant ones (e.g. noise).

We intend to apply the KLT to signals to what are presumed to be a deterministic data. The purpose is also to separate the significant information from possibly added noise but mainly to extract from a huge set of signals its dominant components.

Given a set of signals

$$u(n, x) \quad \text{with} \quad x \in [a, b], \tag{4}$$

where x is the distance and n the number of the measurement, the KLT provides a decomposition

$$u(n, x) = \sum_{i=1}^{\infty} a_i(n) \, \psi_i(x). \tag{5}$$

The *characteristic functions* (CF) are denoted by, $\psi_i(x)$, representing basic patterns and the *amplitudes*, denoted by $a_i(n)$, are their contribution in each signal.

This resembles the well known Fourier transform (FT), where a signal is described as a superposition of basic functions ($\sin(x)$ resp. $\cos(x)$) multiplied by certain amplitudes. The important difference between KLT and FT is that the basic functions of the FT are fixed, whereas the characteristic functions of the KLT are determined by and thus are specific to the set of signals $u(n, x)$. The two basic requirements of the KLT are first, the characteristic functions $\psi_i(x)$ form an orthogonal basis:

$$\langle \psi_i(x), \psi_j(x) \rangle = \int_a^b \psi_i(x) \, \psi_j(x) dx = \delta_{ij}. \tag{6}$$

Second, the amplitude functions $a_i(n)$, which are the projections of the signals onto this base,

$$a_i(n) = \langle u(n, x), \psi_i(x) \rangle = \int_a^b u(n, x) \, \psi_i(x) dx \tag{7}$$

are uncorrelated

$$\boldsymbol{E}_n(a_i(n) \, a_j(n)) = \lambda_i \delta_{ij}, \tag{8}$$

where $\boldsymbol{E}_n$ is the expected value over n.

These requirements are met by $\psi_i(x)$, which are the eigenfunctions of the integral operator

$$\int_a^b r_{uu}(x', x) \, \psi_j(x) dx = \lambda_j \psi_j(x') \tag{9}$$

where $r_{uu}(x', x)$ is the autocorrelation function of the signals

$$r_{uu}(x', x) = \boldsymbol{E}_n(u(n, x') \, u(n, x)). \tag{10}$$

Next, we consider the approximation

$$u_N(n, x) = \sum_{i=1}^{N} a_i(n) \, \psi_i(x). \tag{11}$$

The *energy* contents covered by the approximation $u_N(n, x)$ is

$$\boldsymbol{E}_n(u(n, x)\, u(n, x)) = \sum_{i=1}^{N} \boldsymbol{E}_n(a_i^2(x)) = \sum_{i=1}^{N} \lambda_i. \tag{12}$$

Thus, if one chooses the characteristic functions ψ_i belonging to the N largest λ_i for the approximation then it covers the highest possible information content of the set of signals, and the mean square approximation error is as small as possible.

The efficiency of the KLT arises from the fact that it is *optimal* in the sense that for all N the approximation with the first N characteristic functions covers the most information contents of the signals (or the mean square approximation error is a minimum) amongst all possible other bases $\phi_i(x), i = 1, ...N$. We employ this optimality feature for the purpose of best information compression.

All this carries quite directly over to the *discrete* Karhunen–Loève–Transformation which we will apply, because every measurement naturally is a discrete set of data, where

$$u(n, x_i) = \boldsymbol{u}(n) \tag{13}$$

with $i = 1, ..., d$ and $\boldsymbol{u} \in \mathbf{R}^d$ and the characteristic functions, $\boldsymbol{\psi}_i \in \mathbf{R}^d$, are the eigenvectors of the covariance matrix

$$(\boldsymbol{C}_{uu} - \lambda_i \boldsymbol{I})\, \boldsymbol{\psi}_i = 0 \tag{14}$$

with

$$\boldsymbol{C}_{uu} = \boldsymbol{E}_n(\boldsymbol{u}(n)\, \boldsymbol{u}^T(n)). \tag{15}$$

The number of measurements on the real railway track was too small to allow for a thorough study of the KLT procedure. Therefore, measurements had to be produced on laboratory level.

5   Small Scale Experiment

We performed small scale laboratory experiments . We installed angular velocity sensors in the coach in order to collect information from vehicle behavior for different track conditions. The track is 10m long and allows changes of the height as well as of the sleepers.

We present measurements along a 2 m test section. The sleepers are spaced at 20 cm and we varied the height of the sleepers no. 6 and 7 at 1 m and 1.2 m.

The measured response is of course a dynamic reaction of the system to the impact of track failure. This is also obvious by comparison of the responses of the journey there, Fig. 6 and the return journey, Fig. 7, to the same impact.

The information contained in this set of signals has to be compressed and evaluated efficiently. We apply the Karhunen–Loève-Transformation in order to detect the basic components of this set of signals and their contribution to every single signal.

Often the KLT is applied to the mean value free signal $E_n(\boldsymbol{u}(n)) = 0$ in order to pronounce the actual variations in the set. We are especially interested, however, in variations with respect to the initial signal. Therefore, we apply the KLT to the difference signal with respect to the first signal:

$$\boldsymbol{u}(n) = \dot{\boldsymbol{\alpha}}(n) - \dot{\boldsymbol{\alpha}}(1). \tag{16}$$

First we calculate the covariance matrix, $\boldsymbol{C}_{uu}$, of the signal set. The eigenvectors of this covariance matrix are the basic components of the set $u(n, x)$; the eigenvalues corresponding to the eigenvectors are a measure of their significance (energy contribution).

The KLT reveals the major features of what we see in Fig. 6. The set of signals seems to consist of only one basic pattern, which is contained in each signal, with different 'amplitudes'. Analysis of the eigenvalues of $\boldsymbol{C}_{uu}$ confirms that only one eigenvector is significant. This eigenvector covers already 96% of the energy of the set. The second resp. third eigenvector covers merely 3.4% resp. 0.6% of the energy. Figure 8 shows the first eigenvector, i.e., the first characteristic function, $\boldsymbol{\psi}_1$. Thus, the whole set of signals is well approximated by the first eigenvector, and its information compressed to a single characteristic function and its amplitude function.

The height of the sleepers no. 6 and 7 was decreased until journey 10 and then it was increased again. We conducted two measurement journeys per height. Figure 9 shows the first amplitude function, i.e., the contribution of the first CF, $\boldsymbol{\psi}_1$, to each signal.
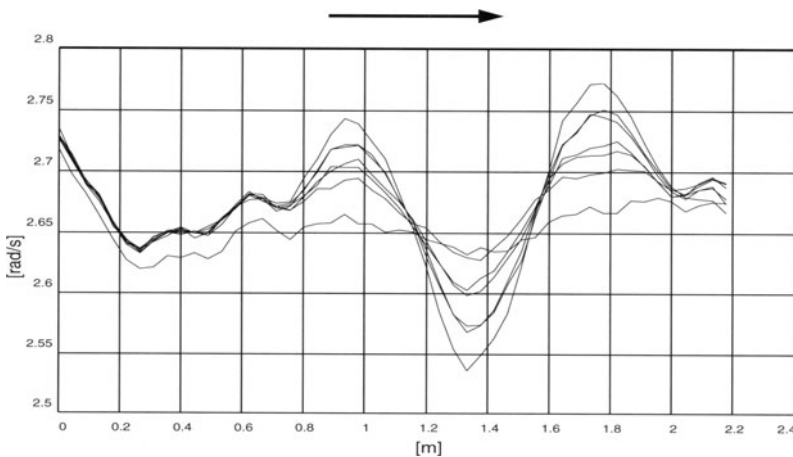
Fig. 6. Measured angular velocity (pitch) of the coach under varying track conditions. The arrow indicates the direction of the train velocity along the track
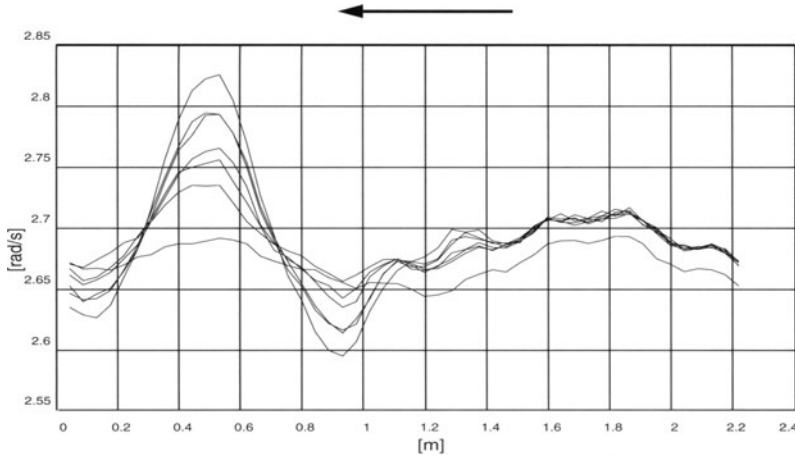
Fig. 7. Measured angular velocity (pitch) of the coach under varying track conditions, return journey
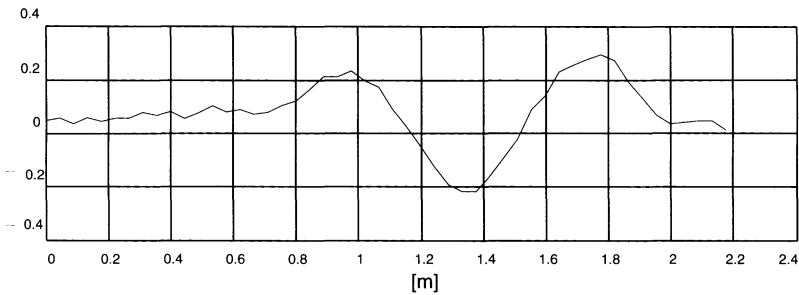
Fig. 8. First characteristic function ψ_1 of the set of signals in Fig. 6
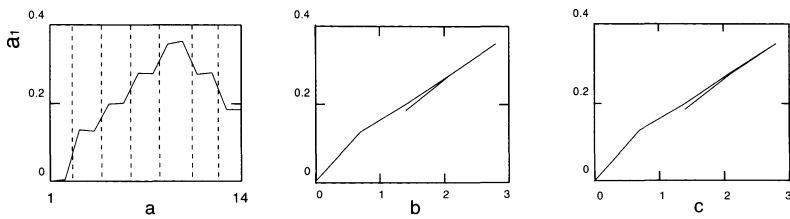
Fig. 9. First amplitude function: **(a)** $a_1(n)$ contribution of the first CF to each journey there, **(b)** a_1 as a function of the settling of the sleepers in mm, **(c)** same as **(b)** but for the return journey

The first amplitude function reveals: First, it is nearly identical for identical track conditions (for the two journeys per height). Second, a_1 is nearly independent of the history of the height of the sleepers, i.e., there is no hysteresis in Fig. 9 b and c. Third, a_1 is nearly a *linear function* of the settling of

the sleepers in mm. Fourth, although the first characteristic function of the return journey is completely different from the one of the journey there, its amplitude reflects the severeness of the failure in the same way.

The application of the KLT to the data of our small scale laboratory experiment shows that the Karhunen–Loève-Transformation is an efficient tool for information compression. It is also capable of displaying the severity of a track failure.

6   Evaluation of Full Scale Measurements

Measured data from real tracks look different than those from small scale experiments for several reasons:

- The vehicle dynamics is influenced by primary and secondary suspension.
- The railway track is a complicated structural component with respect to geometry and dynamics.
- Weather and climate.

Data from three journeys already span a three dimensional manifold, whose information is not compressed by three CFs resp. one reference signal and two CFs in the strict sense. The variations within one signal ensemble are relatively large compared to the expected changes of the track within a short measurement periods. Therefore it seems that not the variations in the signal ensemble (difference e.g. to the mean of all data) contain important data but just the mean of the ensemble at best. If not the changes are to be approximated in an optimal sense, but the entire information of the signal ensemble, only the original data can be used. In general, it can be assumed that more measurements produce a larger data ensemble, which spans a higher dimensional manifold. To the same amount more CFs are needed for a good approximation. But, more CFs give more information about the deviations of the signal for almost the same track rather than about the track itself. The question, whether the KLT can produce significant information beyond the simple mean of all data from subsequent journeys, can not yet be answered. Even the assumption that the dynamic reactions of the train on a prominent excitation are rather clear and dominant, cannot be certified in all cases.

This can be exemplified with measured data obtained on the new high speed track between Hannover and Göttingen. Figure 10a depicts the angular velocity of the pitch angle between stations km78.25 and km78.55 of three journeys. The same signal, but low pass filtered (5Hz LP) is shown in Fig. 11a. At station km78 a curve begins with a ramp of the outer rail. This ramp ends at station km78.3 . Obviously the ramp causes vibrations of the car body. But as synchronous the oscillations are at all three journeys as obvious the differences are between the three journeys. In addition, these differences are larger than the differences at the 'normal' part of the track before the ramp. Figures 10b and 10d show the first and the second CF of the original

(a)

(b)   (c)

(d)   (e)
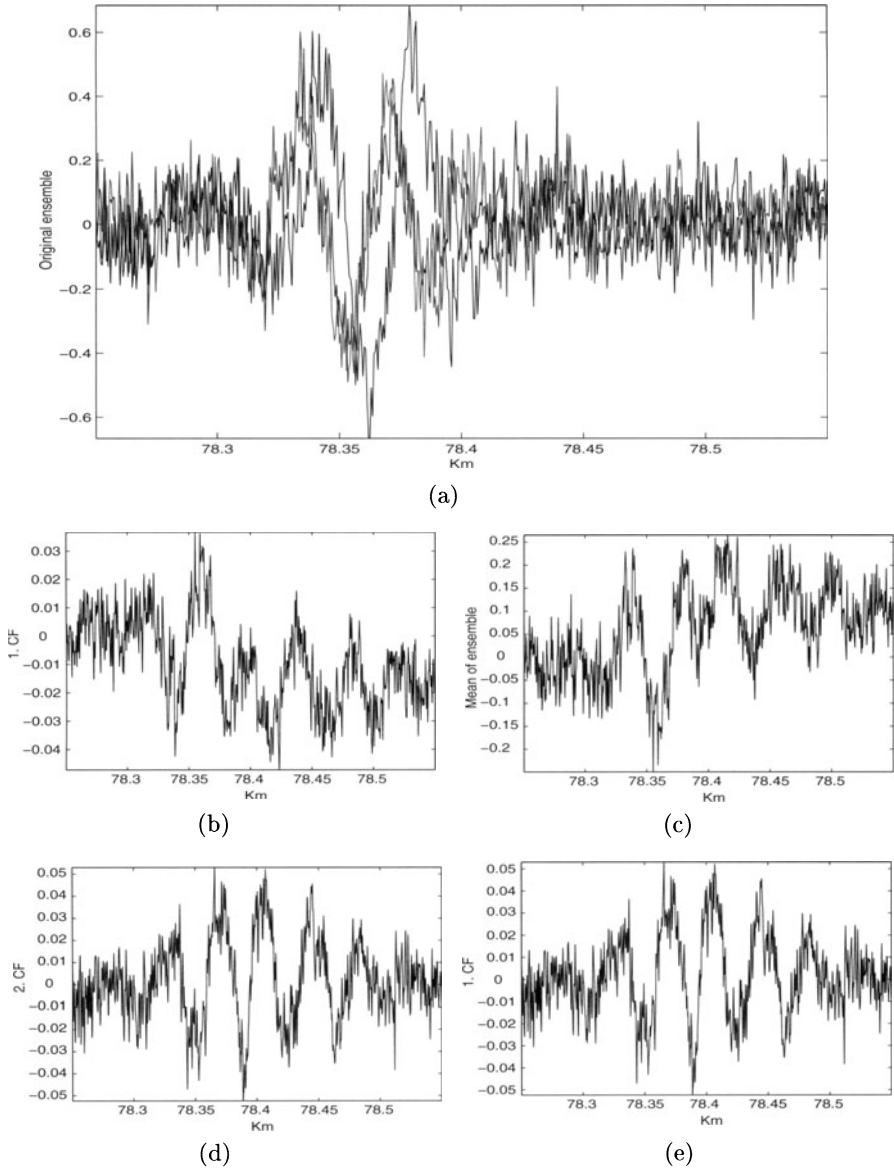
Fig. 10. KL-Transformation of three angular velocities of the pitch angle. (**a**) Entire ensemble. (**b**), (**d**) First and second CF of the original data. (**c**) Mean of all data. (**e**) First CF of the data after subtraction of the mean value

signals. For comparison the mean of the data (Fig. 10c) and the first CF of the signal with the mean value subtracted (Fig. 10e) is placed to the right of the corresponding plots. Here the first CF of the KLT of the measured signals essentially corresponds to the mean of all signals (see e.g. Fig. 11c).

Apart from signs, which are not fixed at characteristic functions because of
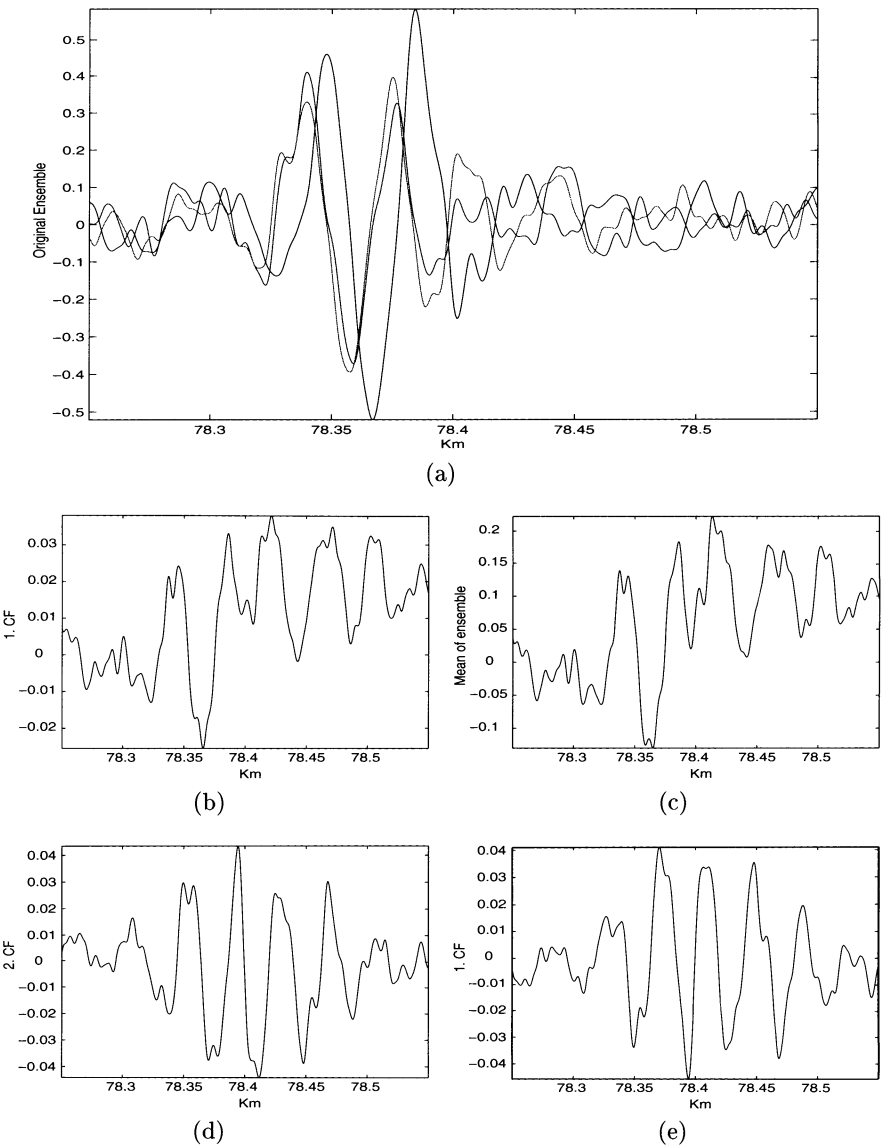
Fig. 11. Same as Fig. 10, but the original data filtered (5 Hz LP)

the similarity of characteristic functions to eigenfunctions, and from scaling the adjacent curves are in good agreement. This implies that the most important information (the first CF of the original signal) is more or less

equivalent to the mean of all data of subsequent passages. In this very case the computation of the KLT is not necessary, strictly speaking.

7   Evaluation of the Results and Outlook

In principle, measurements obtained by our measurement platform on regularly scheduled trains are technically feasible. For the purposes of anti-skid, pulses from a free running axle are normally processed, anyway, and the *Indusi*–signals of the track monitoring system are registered as well, e.g., for the automatic control of trains when signals are overlooked. At any rate, the transmission of these signals to the measurement platform means a certain technical effort.

In order to evaluate if such an effort is worth-while, apart from the explanations already mentioned the following has to be pointed out:

By experimental means was proved that the reactions of the vehicle depend substantially on the vehicle itself (non-roundness of the wheel etc.), i.e. the reactions do only depend on the track. When looking for alterations in the signals of the vehicles, e.g., by means of the KL-Transformation, the question arises whether these changes result from alterations of the vehicle or of the track. Or put in different words: What has worn more rapidly, the vehicle or the track? Is it possible to distinguish these effects, e.g., by filtering? This, of course, assumes linear superposition of the effects.

From the actual state of the art this question cannot be answered. One proposal is to find out by as many measurements as are statistically significant if the effect of outliers can be eliminated.

One indicator of an 'interesting' position can be the variance of the ensemble of consecutive rides as function of the position. The kind of the dynamic reaction can be observed at best by the mean value of all rides. But all significant reactions observed up to now were identical, though: vibrations with a frequency of roundabout 1 Hz.

Whether the kind of defect tracks or even a slow development of defects can be detected from these signals, cannot be answered without further detailed tests and evaluation of the large amount of measurements.

8   Conclusion

This article describes an approach for a continuous track survey in the Deutsche Bahn AG railway network. With this approach a systematic and comparative inspection and assessment of railway tracks will be possible. Instead of expensive special measurement trains, we need only relatively inexpensive equipment installed in the coach of a wagon.

The first idea is to detect the track settlement due to accumulated dynamic loads by means of an inverse dynamic system that reconstructs the original

input, the track features, from measurements in the coach. Due to the input-quadratic character of the system, however, this approach is not applicable. We present real world measurements. They indicate that the dynamic vehicle-track system has unique asymptotic behaviour, such that we can expect variations of the vehicle response to stem actually from variations of the track or the vehicle.

It is understood that the dynamic response of the vehicle, i.e. the measured signals, are influenced by track irregularities and vehicle features like worn wheels and damaged bogies. It is our belief that the variation of the vehicle response contains significant information about the variation of the vehicle-track system. The specific response (e.g., a certain periodicity for worn wheels) should allow to separate different kinds of failures from each other.

In order to handle the enormous amount of data obtained from the measurements we apply Karhunen–Loève–transformation (KLT) for data compression. The successful application of KLT to a small scale experiment shows that efficient information compression is possible and it enables us to measure the severeness of a failure.

9   Acknowledgment

References

1. Feldmann, U., Hasler, M., and Schwarz, W. (1996) Communication by chaotic signals: the inverse system approach. Int J Circuit Theory and Applications **24**(5), 551–579
2. Fröhling, R. (1997) Deterioration of railway track due to dynamic vehicle loading and spatially varying track stiffness. Ph.D. thesis, University of Pretoria, Pretoria
3. Holmes, P., Lumley, J., and Berkooz, G. (1996) Turbulence, coherent structures, dynamical systems and symmetry. Cambridge University Press, Cambridge
4. Kreuzer, E. and Kust, O. (1996) Analysis of long torsional strings by proper orthogonal decomposition. Archive of Applied Mechanics **67**, 68–80
5. Vidyasagar, M. (1993) Nonlinear Systems Analysis. 2nd edn. Prentice-Hall, Englewood Cliffs N. J.

Combined Modelling of Discretely Supported Track Models and Subgrade Models – Vertical and Lateral Dynamics

Ulf Gerstberger[1], Klaus Knothe[1], and Yongfang Wu[2]

[1] Technische Universität Berlin, Institut für Luft- und Raumfahrt,
   Marchstraße 12, 10587 Berlin, Germany
[2] Shenzhen Metro Co. Ltd.,
   9/F., No. 4009 Shennan Road, 518026 Futian District, Shenzhen, China

Abstract. In the present paper, a frequency domain model for vertical and lateral dynamics and two time domain models for vertical dynamics of the ballasted track are described. All models take into account the subgrade behaviour using halfspace theory. The frequency domain model is discussed in detail with respect to the algorithm for infinite periodic structures which are used for the derivation of the governing equations. As regards the time domain models, emphasis is placed on the particular capabilities to simulate non-linear behaviour. For each model presented in the paper, the results of numerical analyses dealing with vehicle/track interaction are given.

1   Introduction

Objectives. A number of models have been published in literature dealing with the dynamic behaviour of ballasted tracks in the frequency or time domain, some also including subgrade models. However, a more sophisticated accounting for the subgrade often leads to time-expansive algorithms, especially in the time domain. In the following text, models are presented which allow for simulations of vehicle track interaction including the underlying halfspace on the one hand, and using comparatively fast algorithms on the other hand. The main focus is the determination of loads within the entire system of vehicle track and subgrade at high speed traffic. The models may provide tools for investigations concerning long-term behaviour, life-cycle costs and safety issues.

Mechanism of Excitation. The relevant mechanisms of excitation comprise out-of-round wheels, rail-surface irregularities and parametric excitation. These are considered in both frequency and time domain while the impact of voided sleepers on the dynamic loads within the track is dealt with in time domain, only. The frequency domain analysis is restricted to the mid-frequency range up to 600 Hz which is the relevant range for the mechanisms of excitation considered. The two time-domain models presented in the article are capable of simulating non-linear behaviour of the track, such as lift-off of the wheel from the rail or non-linear material behaviour.

2   Frequency Domain Model for Vertical and Lateral Dynamics

Algorithm for Infinite Periodic Structures. In the frequency domain the ballasted track is considered as an infinite, periodic structure placed on an elastic halfspace. The smallest unit of the periodic structure comprises the section of the rails between two sleepers, the discrete support of the rail and the underlying halfspace. The characteristic length of the system is given by the sleeper distance l_s.

Using the eigensolution of the state vector of an arbitrary unit of the model it is possible to define boundary conditions of the infinite structure of the track. Subsequently, the infinite structure is excited by harmonically oscillating loads in vertical and lateral direction.

The loads are applied to the rail-head of the discretely supported rail at an arbitrary position along the track. As the track's geometry is assumed to be symmetrical with respect to the longitudinal axis of the track, any loading condition can be split up into a symmetric and antimetric component.

Model of the Discrete Support. Cross-sections of the track-model are shown in Fig. 1. In the cross section on the left hand side, the degrees of freedom of the rail, the support structure and the subgrade are given. It should be noted, that in the case of the subgrade, the degrees of freedom represent loads acting on the surface of the subgrade. In the cross section depicted on the righthand side of Fig. 1, the load vectors R_i and T_i acting on the support structure are shown.

The support structure comprises the rail-pads, the sleeper, the sleeper-pads and the ballast. The pad is modelled by an elastic element with visco-elastic or structural damping. Due to the frequency dependence of the damping of common rail-pad materials, the use of structural damping may be preferred [3].

The sleeper is modelled as a rigid mass with additional elastic modal degrees of freedom to account for the first symmetric and antimetric bending mode of the sleeper. Further elastic modes may be accommodated in the model if necessary. No rotation of the sleeper around the vertical axis of the track is allowed.

The ballast is modelled using discrete blocks beneath the area of contact of the sleeper and the ballast layer. The mid-span section of the sleeper is assumed to be a no-contact area. In vertical direction the cross-sectional area of the block varies in accordance with the angle of load-distribution within the ballast material. Each block of ballast is modelled as a short rod in vertical direction and as a short section of a beam with respect to the lateral direction and the rotation around the longitudinal axis. For the mechanical model a FE-formulation is used, with each block discretized by one finite element.
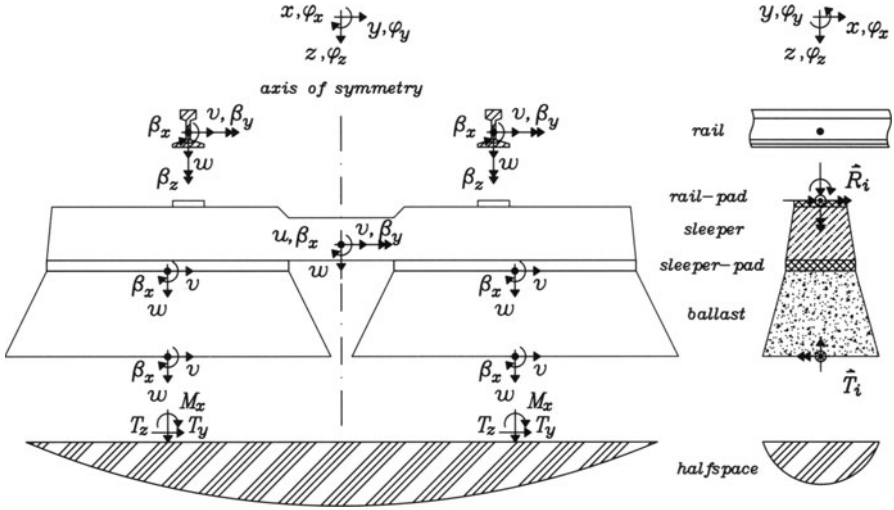
Fig. 1. Cross-sections of the model of the infinite, periodic track with degrees of freedom (*left*) and the loads acting on the discrete support i (*right*)

It is assumed that the ballast block does not exhibit rotations around the lateral axis and no translations along the longitudinal axis of the track.

Due to the harmonic excitation of the track, all forces and displacements oscillate harmonically as well. Thus, time and spatial dependency of the forces and displacements can be separated:

$$F(x,y,z,t) = \hat{F}(x,y,z)e^{i\Omega t} \ \text{ and } \ u(x,y,z,t) = \hat{u}(x,y,z)e^{i\Omega t}. \tag{1}$$

Considering the support structure in Fig. 1 only, it can be seen that "external" loads act on the pad ($\boldsymbol{R}_i$) and the ballast block ($\boldsymbol{T}_i$). Therefore, internal degrees of freedom can be eliminated, leading to the following composition of the vector of unknowns of the support structure i:

$$\hat{\boldsymbol{u}}_{r,i}^{T} = \{\hat{v}_{r,i}, \hat{w}_r, \hat{\beta}_{r,x}, \hat{\beta}_{r,y}, \hat{\beta}_{r,z}\} \ \text{ and } \ \hat{\boldsymbol{u}}_{b,i}^{T} = \{\hat{v}_{bb}, \hat{w}_{bb}, \hat{\beta}_{bb,x}\}, \tag{2}$$

where $\hat{\boldsymbol{u}}_{r,i}$ is the vector of degrees of freedom of the rail and $\hat{\boldsymbol{u}}_{b,i}$ is the vector of degrees of freedom of the ballast block in the area of interaction with the subgrade. Application of the principle of virtual work yields the following set of complex algebraic equations for the support structure:

$$\begin{bmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} \\ \mathbf{K}_{12}^{T} & \mathbf{K}_{22} \end{bmatrix} \begin{Bmatrix} \hat{\boldsymbol{u}}_{r,i} \\ \hat{\boldsymbol{u}}_{b,i} \end{Bmatrix} = \begin{Bmatrix} \hat{\boldsymbol{R}}_i \\ -\hat{\boldsymbol{T}}_i \end{Bmatrix}. \tag{3}$$

The vectors $\boldsymbol{R}_i$ and $\boldsymbol{T}_i$ contain six loads acting in the rail-pad and three loads acting on the ballast block, respectively.

Model of the Rail. With respect to both vertical and lateral vibrations, the rail is modelled by Timoshenko-beams with shearing and bending stiffness. In longitudinal direction the rail is modelled as a torsional rod. No translation in longitudinal direction is allowed. The discretely supported, infinite beam is described by a set of two coupled 2nd order differential equations given below:

$$EI\frac{\partial^2 \beta_r(x,t)}{\partial x^2} - GA_s\{\beta_r(x,t) + \frac{\partial w_r(x,t)}{\partial x}\} - \hat{\mu}_r\frac{\partial^2 \beta_r(x,t)}{\partial t^2}$$

$$= \sum_{-\infty}^{+\infty} \delta(x - jl_s)\, k_p(1 + i\eta)\Delta\beta(jl_s,t) \qquad (4)$$

$$GA_s\{\frac{\partial \beta_r(x,t)}{\partial x} + \frac{\partial^2 w_r(x,t)}{\partial x^2}\} - \mu_r\frac{\partial^2 w_r(x,t)}{\partial t^2}$$

$$= -Q(t)\delta(x) + \sum_{-\infty}^{+\infty} \delta(x - jl_s)\, k_p(1 + i\eta)\Delta w(jl_s,t)\,, \qquad (5)$$

where EI and GA_s is the bending and shearing stiffness and $w_r(x,t)$ and $\beta_r(x,t)$ denote the displacement and the rotation of the cross-section of the beam. Indices denoting the direction of displacement and the axis of bending have been omitted, as the set of equations is applied for both vertical and lateral vibration of the beam.

In the case of torsion, the infinite rod with discrete support is described by one 2nd order differential equation:

$$GI_T\frac{\partial^2 \beta_{r,x}(x,t)}{\partial x^2} - \hat{\mu}_{r,xx}\frac{\partial^2 \beta_{r,x}(x,t)}{\partial t^2}$$

$$= -M_x(t)\delta(x) + \sum_{-\infty}^{+\infty} \delta(x - jl_s)\, k_{p,xx}(1 + i\eta)\Delta\beta(jl_s,t)\,, \qquad (6)$$

where GI_T and $\mu_{r,xx}$ are the torsional stiffness and mass moment of inertia of the rail, respectively.

Shape functions for finite element discretization satisfying the differential equations are given by Koloušek [7] – due to inertia effects, they are frequency dependent. Since the inertia effects of the mass of the rail are taken into account by the shape functions, the section of the rail between two supports is discretized by one finite element only. At the location of each support, five degrees of freedom of the rail exist:

$$\hat{\boldsymbol{u}}_{r,i}^T = \{\hat{v}_r, \hat{w}_r, \hat{\beta}_{r,x}, \hat{\beta}_{r,y}, \hat{\beta}_{r,z}\}_i\,. \qquad (7)$$

Obviously, the model of the rail does not account for independent distortions of the rail head and foot – mode shapes of the rail which become more important in the higher frequency range beyond 600 Hz.

By introducing the harmonic expressions given in (1) for the forces and displacements, the algebraic set of equations for the unloaded section of the rail between two supports yields the following expression in matrix notation:

$$\begin{bmatrix} S_{11} & S_{12} \\ S_{12}^T & S_{22} \end{bmatrix} \left\{ \begin{array}{c} \hat{u}_{r,i} \\ \hat{u}_{r,i+1} \end{array} \right\} = \left\{ \begin{array}{c} \hat{Q}_i \\ -\hat{Q}_{i+1} \end{array} \right\}, \tag{8}$$

where the vectors $\hat{Q}_i$ and $\hat{Q}_{i+1}$ comprise the cross-sectional forces and moments of the rail.

Formulation for the Subgrade. At the centre-point of the contact-area of the ballast block and the subgrade, two forces and one moment are transmitted onto the subgrade. A relation between the loads in vector $\hat{T}_i$ and corresponding displacements and rotations $\hat{u}_{f,i}$ of the subgrade is given by the halfspace theory. The halfspace may either be homogeneous or horizontally layered and the dynamic behaviour is described by three material parameters: the density (ρ), the shear-wave velocity (c_s) and the Poisson's ratio (ν).
In general, the receptance is a scalar value defined by the fraction of the values of the displacement (or rotation) of the surface of the halfspace and the load. The receptance $H_{i,j}$ determined at the location i due to a unit load at location j is given by:

$$H_{i,j} = \frac{u_{f,i}}{T_{f,j}} \quad \left[\frac{\mathrm{m}}{\mathrm{N}} \right]. \tag{9}$$

With respect to the periodic structure of the track, the case $i = j$ yields the direct receptance $H_{0,0}$, whereas the cases $i \neq j$ yield the transfer receptances $H_{0,j}$.
Applying a receptance formulation for the subgrade of the track, scalar values of displacement and rotation $\hat{u}_{f,i}$ in the area of contact of any ballast block i can be expressed in terms of the loads acting in the area itself and in the area of contact $i \pm j$ of the adjacent blocks of ballast and the halfspace:

$$\hat{u}_{f,i} = H_{0,0}\hat{T}_i + \sum_{j=1}^{M} H_{0,j} \left\{ \hat{T}_{i+j} + \hat{T}_{i-j} \right\}, \tag{10}$$

where M is the number of transfer receptances taken into account. Due to the periodicity of the track model a single set of direct and transfer receptances is valid for the receptance formulation of each support structure of the track model.
For reasons of simplification, the matrices of the receptance formulation have elements on the diagonal only. This corresponds to relaxed boundary conditions in the area of interaction of ballast and subgrade. However, the application of non-relaxed boundary conditions, i.e. the incorporation of the complete matrix, is also possible.

$$\mathbf{H}_{0,j} = \begin{bmatrix} H_{0,j}^y & 0 & 0 \\ 0 & H_{0,j}^z & 0 \\ 0 & 0 & H_{0,j}^{\varphi x} \end{bmatrix} \tag{11}$$
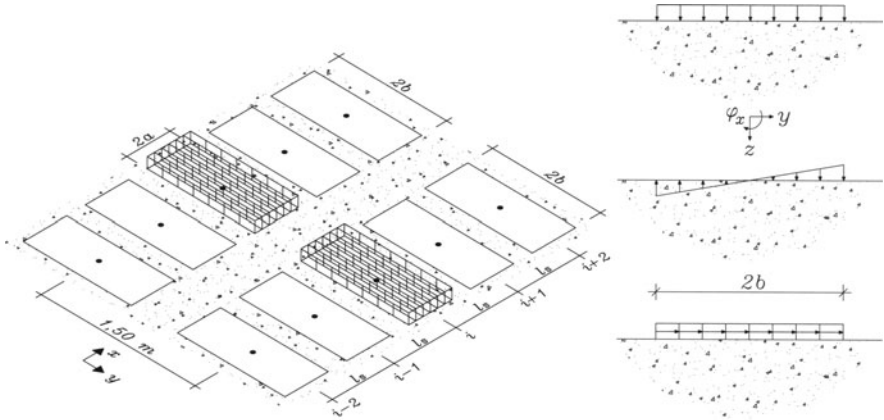
Fig. 2. Areas of interaction of the ballast blocks and the subgrade (*left*) and load distributions of unit loads (*right*) for the determination of the subgrade's receptances

Compatibility between the loads and the degrees of freedom of the ballast block $\hat{T}_i$ and $\hat{u}_{b,i}$ on the one hand, and the loads and distortion of the surface of the halfspace $\hat{T}_i$ and $\hat{u}_{f,i}$ on the other hand, needs to be accomplished. Therefore, the two forces and one moment acting on the ballast block are transformed into equivalent constant and linear pressure-distributions acting in the areas of contact of ballast blocks and subgrade on the surface of the halfspace as shown in Fig. 2. As can be seen on the righthand side in the figure, the vertical and lateral forces are represented by constant pressure distributions whereas the moment is transformed into a linear pressure distribution in vertical direction.

Compatibility of the degrees of freedom of the ballast block and of the halfspace, however, is achieved by determining mean values of the displacements and the rotation of the halfspace in the areas of contact. The mean values $\hat{u}_{f,i}$ are then assigned to the centre-points of the areas of contact.

Eigensolution of the Infinite Track Model. The units of the periodical structure of the track are coupled via the rail as well as via the subgrade. Within the layer of ballast coupling is assumed to be negligible and is therefore not taken into account. The infinite model of the super-structure of the

track is put together by attaching the rail sections at the location of the discrete supports. At any support i the equilibrium of the forces and moments reads:

$$\hat{Q}_{i,L} - \hat{Q}_{i,R} + \hat{R}_i = 0 \,, \tag{12}$$

where $\hat{Q}_{i,L}$ and $\hat{Q}_{i,R}$ are the cross-sectional forces of the rail sections given in (8) and $\hat{R}_i$ denotes the loads due to the discrete support in (3). Considering the unloaded track, the procedure yields an infinite set of algebraic equations for the unknown vectors $\hat{u}_{r,i}$ and $\hat{u}_{b,i}$:

$$\begin{bmatrix} \ddots & \ddots & \ddots & \\ \ddots & \mathbf{A} & \mathbf{B} & \mathbf{0} & \\ \ddots & \mathbf{B}^T & \mathbf{A} & \mathbf{B} & \ddots \\ & \mathbf{0} & \mathbf{B}^T & \mathbf{A} & \ddots \\ & & \ddots & \ddots & \ddots \end{bmatrix} \left\{ \begin{array}{c} \vdots \\ \left\{ \begin{array}{c} \hat{u}_r \\ \hat{u}_b \end{array} \right\}_{i-1} \\ \left\{ \begin{array}{c} \hat{u}_r \\ \hat{u}_b \end{array} \right\}_{i} \\ \left\{ \begin{array}{c} \hat{u}_r \\ \hat{u}_b \end{array} \right\}_{i+1} \\ \vdots \end{array} \right\} = \left\{ \begin{array}{c} \vdots \\ \left\{ \begin{array}{c} 0 \\ -\hat{T} \end{array} \right\}_{i-1} \\ \left\{ \begin{array}{c} 0 \\ -\hat{T} \end{array} \right\}_{i} \\ \left\{ \begin{array}{c} 0 \\ -\hat{T} \end{array} \right\}_{i+1} \\ \vdots \end{array} \right\} \,. \tag{13}$$

The righthand side of (13) comprises the loads $\hat{T}_i$ acting between the blocks of ballast and the subgrade. The loads, however, are unknown values and therefore are introduced into the vector on the left hand side of the equation. Subsequently, the receptance formulation (10) is incorporated into the infinite system of equations by substituting $\hat{u}_{f,i}$ in (10), taking into account the compatibility of the degrees of freedom of the ballast and the subgrade:

$$\hat{u}_{b,i} = \hat{u}_{f,i} \,. \tag{14}$$

Finally, the degrees of freedom of the ballast $\hat{u}_{b,i}$ are eliminated and the infinite system of equations yields:

$$\begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots \\ \ddots & \mathbf{D}_1^T & \mathbf{C} & \mathbf{D}_1 & \mathbf{D}_2 & \mathbf{0} \\ \ddots & \mathbf{D}_2 & \mathbf{D}_1^T & \mathbf{C} & \mathbf{D}_1 & \mathbf{D}_2 & \ddots \\ & \mathbf{0} & \mathbf{D}_2 & \mathbf{D}_1^T & \mathbf{C} & \mathbf{D}_1 & \ddots \\ & & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} \left\{ \begin{array}{c} \vdots \\ \left\{ \begin{array}{c} \hat{u}_r \\ \hat{T} \end{array} \right\}_{i-1} \\ \left\{ \begin{array}{c} \hat{u}_r \\ \hat{T} \end{array} \right\}_{i} \\ \left\{ \begin{array}{c} \hat{u}_r \\ \hat{T} \end{array} \right\}_{i+1} \\ \vdots \end{array} \right\} = \mathbf{0} \,, \tag{15}$$

where the number of matrices $\mathbf{D}_2, \ldots \mathbf{D}_M$, i.e. the bandwidth, depends on the number of transfer receptances taken into account in the formulation (10). At each support, the state vector comprises the degrees of freedom of the rail and the loads acting between the ballast block and the subgrade:

$$\boldsymbol{u}_i = \left\{ \hat{\boldsymbol{u}}_r, \hat{\boldsymbol{T}} \right\}_i , \tag{16}$$

as all other degrees of freedom of the support structure have been eliminated. By considering only one hyperrow of the infinite set of algebraic equations of the unloaded track model, for instance the row of vector $\boldsymbol{u}_i$, the problem can be transformed into an eigenvalue problem. The complete eigensolution of the state vector $\boldsymbol{u}_i$ comprises N eigenvectors and eigenvalues. The k-th eigenvector of the eigensolution reads

$$\boldsymbol{\varphi}_k = \psi_{lk} , \quad k, l = 1, \ldots N , \tag{17}$$

where N depends on the number of transfer receptances taken into account in (10).

The eigenvectors of the two adjacent units, $\boldsymbol{\varphi}_{i\pm1,k}$, are characterised by the fact that they only differ from vector $\boldsymbol{\varphi}_k$ by a constant factor, the symplectic eigenvalue λ:

$$\boldsymbol{\varphi}_{k,i-1} = \frac{1}{\lambda} \psi_{lk} \tag{18}$$

$$\boldsymbol{\varphi}_{k,i+1} = \lambda \psi_{lk} . \tag{19}$$

The eigenvalues are determined by introducing the expressions

$$\boldsymbol{u}_{i-1} = \frac{1}{\lambda} \boldsymbol{u}_i \tag{20}$$

$$\boldsymbol{u}_{i+1} = \lambda \boldsymbol{u}_i \tag{21}$$

into the hyperrow i in (15) leading to a system of algebraic equations with coefficients which are functions of the eigenvalues λ and $\frac{1}{\lambda}$. The characteristic polynomial of the eigenvalue problem, a polynomial of the order of N, can be derived as a function in $\left(\lambda + \frac{1}{\lambda}\right)$. Subsequently, the $2 \times N$ eigenvalues can be determined from the complex roots of the polynomial. They occur in sets of two values λ_k and $\bar{\lambda}_k$ where one eigenvalue is the reciprocal of the other. Since the eigenvalues are symplectic, the eigensolution is split up into two parts. The eigensolution of one set of eigenvalues satisfies the boundary conditions for $j \to +\infty$ and the eigensolutions of the other the boundary conditions for $j \to -\infty$.

Having determined the complex eigenvalues, the corresponding eigenvectors can be extracted from the algebraic system. The complete solution of the state vectors, yields two separate expressions,

- for $j > 0$ and $|\lambda_k| < 1$:

$$\boldsymbol{u}_i = \psi_{lk}^0 q_k \;\; \text{and} \;\; \boldsymbol{u}_{i+j} = \lambda_k \psi_{lk}^0 q_k = \psi_{j,lk} q_k \;\; \text{and} \tag{22}$$

- for $j < 0$ and $|\frac{1}{\lambda_k}| < 1$:

$$\boldsymbol{u}_i = \bar{\psi}_{lk}^0 \bar{q}_k \;\; \text{and} \;\; \boldsymbol{u}_{i-|j|} = \frac{1}{\lambda_k} \bar{\psi}_{lk}^0 \bar{q}_k = \bar{\psi}_{|j|,lk} \bar{q}_k \,. \tag{23}$$

Once the complete eigensolution of the state vector $\hat{\boldsymbol{u}}_i$ is determined, it is possible to set up the system of equations of a loaded section of the infinite track. Loads applied to the rail-head at an arbitrary location can be transformed to nodal loads acting on the rail at the location of the supports of the rail section considered.

Using the eigensolution in expression (22) and (23), boundary conditions for the infinite structure extending on both sides of the rail section can be formulated. The solution of the resulting linear set of equations yields the unknown multipliers q_k and $\bar{q}_k$ of the eigensolution from which the solution for all state vectors of the infinite structure can be generated.

Results of Wheel/Track Interaction. By applying a harmonically oscillating load to the rail, the displacements of all components and loads acting within the infinite model can be determined. However, it is often of particular interest to express the dynamic behaviour of the track in terms of the receptance of the rail. The receptance is a function of the complex frequency of excitation $(i\Omega)$ and in the case of vertical excitation it is defined by:

$$H_{rail}(i\Omega) = \frac{\hat{w}_r(x = x_0)}{\hat{Q}_z(x = x_0)} \tag{24}$$

Considering linear behaviour only, the vertical interaction of a wheel-set and the track due to a harmonic out-of-round of wheel with amplitude Δz can be simulated in the frequency domain. Gyroscopic effects as well as the variation of stiffness of the track due to the discrete support are neglected, as the effects on the contact force are minor in comparison to the effect of the out-of-round of wheel. The determination of the vertical dynamic load at the contact of wheel and rail yields a simple formula:

$$\Delta Q_{dyn} = \frac{1}{c_H + \frac{1}{H_{rail}(i\Omega)} + \frac{1}{H_{wheel}(i\Omega)}} \Delta z \,, \tag{25}$$

where c_H, H_{rail} and H_{wheel} are the Hertzian contact stiffness and the receptances of the track and the rigid wheel-set, respectively.

In Figure 3 the variation of vertical dynamic force at the contact of wheel and rail is plotted over the range of frequencies up to 200 Hz. The amplitude
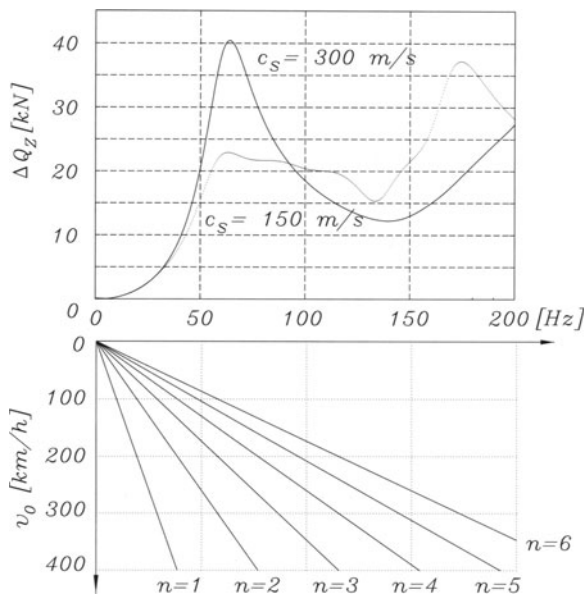
Fig. 3. Vertical dynamic force between rail and wheel due to a harmonic out of round of wheel with amplitude $\Delta z = 0.1\,\mathrm{mm}$ [1]. The two curves represent the cases of a standard track with stiff (*straight line*) and with soft subgrade (*dotted line*)

of out-of-round of wheel is $\Delta z = 0.1\,\mathrm{mm}$ and the parameters of the wheel-set correspond to a $S1002$ wheel-set. The parameters of the track are given in Tab. 1. The two curves depicted in the diagram in Fig. 3 correspond to the results for a stiff and a comparatively soft subgrade, indicated by the shear wave velocity (c_s). Below the two curves the Campbell diagram is given, correlating the velocity of the wheel-set to the corresponding frequency for the harmonics $n = 1, \ldots, 6$.

Table 1. Parameters of the track model of a standard track with B70 concrete sleepers and ZW 700 rail-pads on homogeneous halfspace.

rail-pad [1]		sleeper	ballast				subgrade
k_p [N/mm]	η [−]	m_s [kg]	h_b [m]	ρ_b [kg/m^3]		E_b [N/m^2]	ρ_{sub} [kg/m^3]
150, 0	0.13	284.0	0.3	1800.0		$2.08 10^8$	1800.0

[1] complex dynamic tangent stiffness $k_p(1 + i\eta)$

Using the Campbell diagram, it can easily be deduced that in the case of the stiff subgrade $(c_s = 300\,\mathrm{m/s})$ and a "triangular" out-of-round of wheel $(n = 3)$ the maximum dynamic force occurs at a velocity of approximately $200\,\mathrm{km/h}$. The load is considerably lower in the case of the soft subgrade. The results confirm the dominant influence of the subgrade on the dynamic force in the mid-frequency range.

3 Fast Time Domain Model for Vertical Dynamics

Nonlinear Behaviour. If one is interested in the loads acting between wheel and rail due to out-of-round wheels, the essential range of frequencies is up to 600 Hz. The vertical dynamic behaviour of the track in this range of frequencies is thoroughly described by the receptance of the rail (24). However, the validity of frequency domain models in general is limited, due to the fact that the dynamic loads must not exceed the static loads. In the case of total unloading, lift-off of the wheel from the rail is most likely to occur and non-linear contact formulations have to be used to simulate the interaction of wheel and rail accurately.

The model described in this section will be referred to as the "fast time-domain model". It is based on linear models of the track and the wheel-set but a non-linear contact condition between wheel and rail. The excitation due to out-of-round wheels or rail surface irregularities is incorporated in the formulation of the relative displacement of wheel and rail leading to dynamic vibrations of the entire system.

Time Domain Model of the Track. The transformation of the track model into time-domain is achieved by means of approximation of the track receptances. Based on the approximation function, the track model can be described in terms of ordinary differential equations which are dealt with using standard time-step integration algorithms.
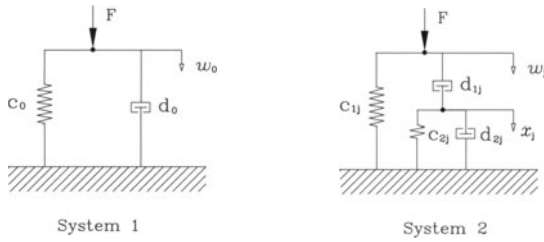
Fig. 4. Mechanical models corresponding to the differential equations (28) (*left*) and (29) (*right*) [9]

For the approximation of the complex receptance of the rail (24), broken-rational functions are used. The order of the denominator polynomial is chosen to be by one greater than the order of the enumerator polynomial to allow for general application of the procedure for track receptances (see also [9]):

$$R_s(i\Omega) = \frac{a_0 + a_1(i\Omega) + a_2(i\Omega)^2 + \cdots + a_{n-1}(i\Omega)^{n-1}}{1 + b_1(i\Omega) + b_2(i\Omega)^2 + \cdots + b_n(i\Omega)^n}. \tag{26}$$

The broken rational function can be split into a sum of partial fractions using partial fraction expansion:

$$R_s(i\Omega) = \frac{\hat{w}_0}{\hat{F}} + \sum_{j=1}^{k} \frac{\hat{w}_j}{\hat{F}} = \frac{1}{c_0 + d_0(i\Omega)} + \sum_{j=1}^{k} \frac{A_j(i\Omega) + B_j}{(i\Omega)^2 + \zeta_j(i\Omega) + \omega_j^2} . \quad (27)$$

Considering the harmonic approach given in (1), each of the partial fractions is interpreted as the receptance of a distinct mechanical model. The differential equations corresponding to the partial fractions are of first and second order:

$$d_0 \frac{dw_0}{dt} + c_0 w_0 = F \quad (28)$$

$$\frac{d^2 w_j}{dt^2} + \zeta_j \frac{dw_j}{dt} + \omega_j^2 w_j = A_j \frac{dF}{dt} + B_j F . \quad (29)$$

The order of the denominator polynomial determines the number of the 2nd order differential equations obtained from the partial fraction expansion of the approximation function. Furthermore, in the case of an uneven number of the order of the polynomial, no 1st order differential equation occurs.

Appropriate mechanical models for the differential equations are shown in Fig. 4. It can be depicted from the figure that $w_0, \ldots, w_j$ are internal degrees of freedom of the models. However, their sum yields the receptance of the rail.

The receptance (24) varies in longitudinal direction of the track depending on the position of the excitation force. Therefore, the parameters of the time-domain model of the track are periodic functions of the longitudinal axis of the track. It should also be noted, that the algorithm is not restricted to numerical data, as receptances obtained from field measurements may also be approximated using broken-rational functions.

Results from Wheel/Track Interaction. In many cases, the influence of the variation of stiffness along the track can be neglected as the influence of parametric excitation on the contact force of wheel and rail is minor in comparison to the influence of out-of-round wheels. In the following example, a track model with constant parameters has been used. The vehicle is modelled as a rigid wheel-set with primary suspension. The wheel-set is fixed to an inertial system which travels at constant velocity along the rail. The gravitational force of the car-body and the bogie is applied at the primary suspension. The contact between wheel and rail is modelled using a non-linear Hertzian contact spring.

In the diagram in Fig. 5 the receptance of the track model and the approximation are depicted. The parameters correspond to standard track with B70 concrete sleepers and a stiff rail-pad [6] on a stiff subgrade ($c_s = 300\,\mathrm{m/s}$). In the figure, the mechanical model of the track derived from the approximation
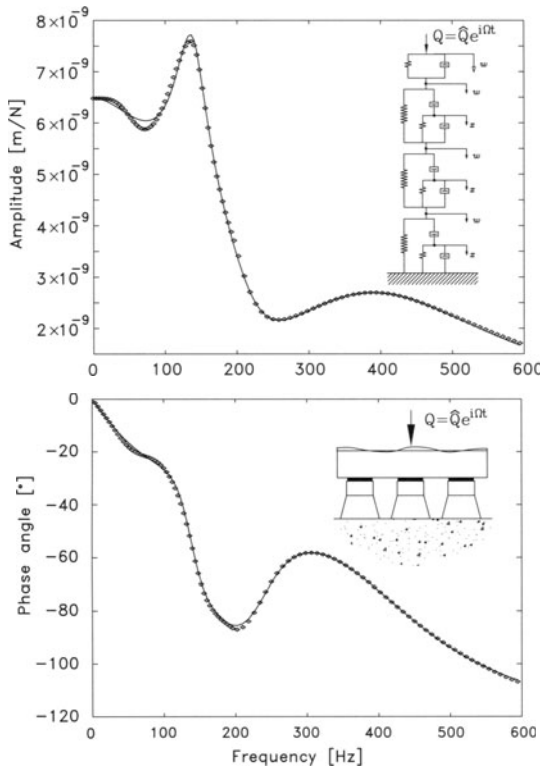
Fig. 5. Absolute value and phase angle of the receptance and the approximation function of a standard ballasted track with B70 concrete sleepers, a stiff rail-pad and a stiff subgrade ($c_s = 300\,\mathrm{m/s}$). The sketch in the top diagram represents the mechanical model of the track derived from the approximation function

is shown as well. As can be seen in the sketch, the model is composed of four elements from the two types shown in Fig. 4.

In Figure 6, results from a non-linear analysis of wheel-track interaction are given in terms of the variation of the vertical contact force between wheel and rail. The system is excited by a harmonic out-of-round of wheel with a "triangular" pattern (n=3). The two curves correspond to the two different amplitudes of out-of-round of wheel considered. It can be seen in the plot, that in the case of the amplitudes of out-of-round of $\Delta z = 0.1$ mm the vertical dynamic load in the wheel/rail contact exceeds the static load and lift-off of the wheel from the rail occurs at regular intervals.

4   General Time Domain Model for Vertical Dynamics

General Non-Linear Behaviour. In order to account for a greater variety of non-linear mechanisms in the superstructure of the track, more complex time-domain models have to be used. In the following, a non-linear time domain model is presented dealing with the vertical interaction of a vehicle travelling at a constant velocity along the ballasted track. The model, referred to as the "general time-domain model", allows for analyses concerning the
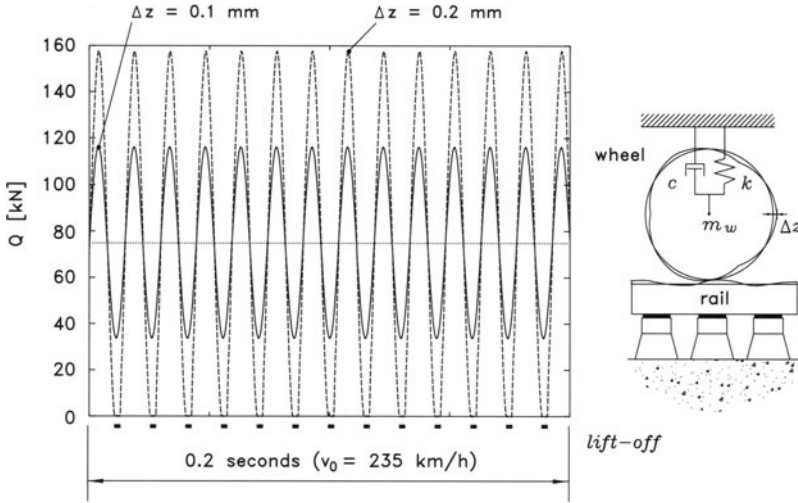
Fig. 6. Variation of vertical force between wheel and rail due to a harmonic out-of-round of wheel (n=3). The amplitude of out-of-round of the wheel is varied between $\Delta z = 0.1\,\mathrm{mm}$ (*straight line*) and $\Delta z = 0.2\,\mathrm{mm}$ (*dotted line*)

effects of out-of-round wheels, the variation of the track stiffness and voided sleepers on the dynamic behaviour [6]. The model is capable of simulating lift-off phenomena of the wheels from the rail, the sleepers from the ballast and also of the ballast from the subgrade. Voided sleepers can be pre-defined at arbitrary positions along the finite track and the incorporation of a non-linear material-law for the ballast is feasible. The latter may be of particular interest regarding investigations concerning ballast deterioration and long-term behaviour.

A sketch of the mechanical model of the track and the vehicle is depicted in Fig. 7. The rail is modelled as a discretely supported Timoshenko-beam with additional elastic modal degrees of freedom using one finite element for the section of the rail between two supports. The structure of the discrete support is similar to that described in section 2. The combined stiffness of the rail fastening and the rail-pad as well as the stiffness of the sleeper-pad are modelled by spring and dash-pot in parallel. However, in contrast to the frequency domain model, the elastic elements assume non-linear visco-elastic behaviour. The contact between ballast block and the subgrade allows for lift-off of the ballast block from the surface of the subgrade. The model of the track is finite and the boundary conditions at both ends of the model are chosen in such a way as to form a ring-model. The number of sleeper-spacings to be taken into account depends on the demand for the influence of revolving waves propagating in the rail to remain negligible. The model of the vehicle is composed of rigid bodies representing two wheel-sets, the bogie and half of

vehicle

φ_d

w_d

w_{w1}

w_{w2}

discretely
supported rail

ballast

subgrade
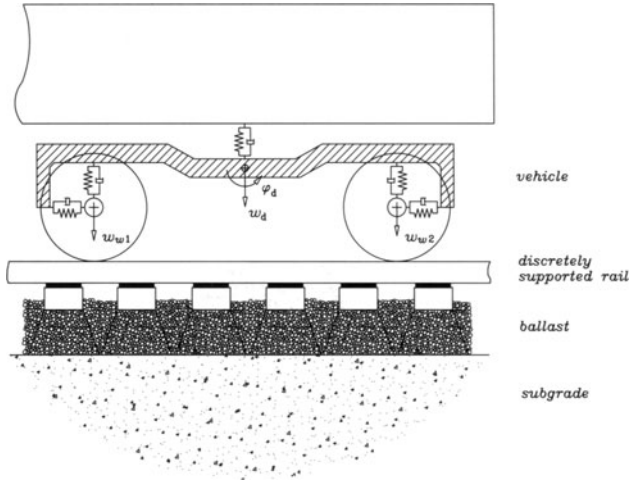
Fig. 7. Sketch of the "general time domain model" for vertical dynamics. The vehicle moves with constant velocity along the ballasted track with discretely supported rail and underlying halfspace

the car body. The vertical plane intersecting the longitudinal axis of the track is the plane of symmetry with respect to the model and the mechanisms of excitation.

Time Domain Model of the Subgrade. The model of the subgrade is based on vertical direct and transfer receptances given in (9) of the halfspace derived from frequency domain analyses. The receptances of the subgrade are approximated using broken-rational functions and subsequently transformed into time domain. The procedure of derivation of the mechanical models is described in detail in section 3 in the case of the receptances of the entire track.

The system of mechanical models corresponding to the direct and transfer receptances of the subgrade are attached "in row" to the degree of freedom of the ballast below each support. In Figure 8 the absolute value of a direct receptance and of the approximation function is plotted over the range of frequencies up to 400 Hz. Furthermore, the mechanical model derived from partial fraction expansion of the approximation function is depicted.

The coupling condition in the receptance formulation (10) of the subgrade determines the total number of mechanical models taken into account. For instance, for a stiff subgrade ($c_s = 300$ m/s) the consideration of two transfer receptance suffices, leading to three subgrade models beneath each ballast block. In the case of softer subgrade materials, a greater number of receptances has to be taken into account, leading to more mechanical models attached to each ballast block. Consequently, the total number of degrees of freedom of the entire system of vehicle, track and subgrade depends on the number of sleepers taken into account and the stiffness of the subgrade.
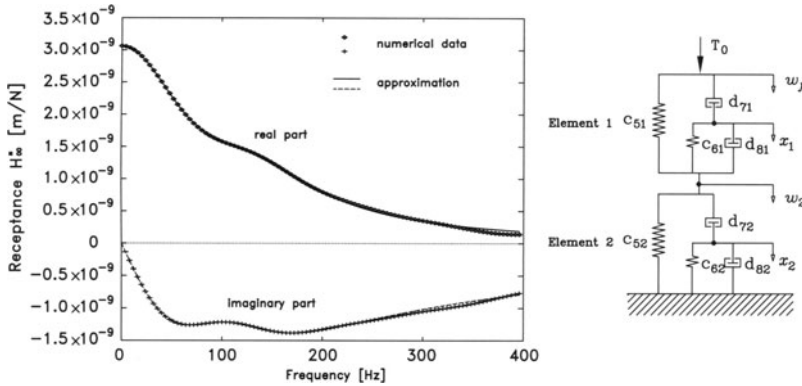
Fig. 8. Direct receptance of a stiff subgrade ($c_s = 300\,\mathrm{m/s}$) and its approximation by a broken rational function. The set mechanical model was derived from the approximation and the corresponding ordinary differential equations can be used for time-step integration

Results from Vehicle/Track Interaction. A vehicle running over a track with voided sleepers inherently induces non-linear behaviour into the mechanical system. The dynamic behaviour of the system depends on the particular combination of out-of-round wheels and void patterns along the track.

In the following example, a standard ballasted track with B70 concrete sleepers, a stiff rail-pad and a stiff subgrade has been used [6]. The mechanical model of the direct receptance of the subgrade is given in Fig. 8. The model of the vehicle corresponds to an ICE passenger car. In the analysis, a track model with 30 sleepers has been considered. In this case, the total number of degrees of freedom of the entire system of vehicle, track and subgrade yields 900.

In Fig. 9 the variation of vertical contact force between the trailing wheel and the rail along a $10\,m$ section of the track is shown. In the plot, three curves are depicted, corresponding to three different combinations of void-patterns and out-of-round of wheels:

1. a vehicle with out-of-round wheels on a track without voided sleepers,
2. a vehicle with ideal round wheels on a track with voided sleepers and
3. a vehicle with out-of-round wheels on a track with voided sleepers.

It can be seen from the plot, that the force yields significantly higher values due to the effect of out-of-round wheels than due to the voided sleepers. In the case of out-of-round wheels on a track with voided sleepers, the void pattern leads to a reduction of peak values of the contact force between wheel and rail.

It should be noted, that in Fig. 9 only the variation of contact force has been considered. The overall loading of the track increases significantly due to the effects of voided sleepers and out-of-round wheels, leading also to a higher loads acting on the ballast see [6].
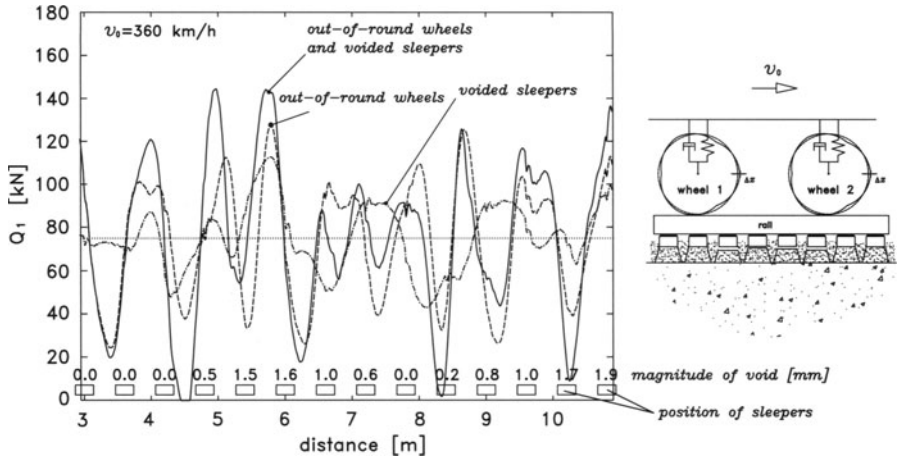
Fig. 9. Variation of vertical dynamic forces acting between rail and wheel due to out-of-round wheels and voided sleepers [6]. The maximum amplitude of out-of-round does not exceed $\Delta z = 0.3\,\mathrm{mm}$, the magnitudes of the voids are depicted in the plot

5   Conclusions

Three models of the ballasted track have been presented in the paper combining models of the superstructure of the track and subgrade models in both frequency and time domain.

In the case of the frequency domain model, the dynamic response in lateral and vertical direction was considered at symmetric and antimetric harmonic excitation of the rail. For the derivation of the model, an algorithm for infinite periodic structures was used. The algorithm allows for

- a more sophisticated definition of boundary conditions of the infinite track,
- detailed modelling of the discrete supports and
- the incorporation of a receptance formulation for the underlying subgrade.

Comprehensive parametric studies may be conducted using the model, as the calculation time is almost negligible. Furthermore, it should be mentioned that the algorithm can also be applied when modelling different types of railway tracks, e.g. for the ballasted track with frame-sleepers.

The time domain models presented in the paper both deal with the vertical dynamics of the ballasted track at symmetric loading. The "fast time domain model" is capable of simulating non-linear contact conditions of the wheel and the rail, allowing for lift-off phenomena to be studied. The track model is based on track receptances which can be either obtained from numerical analyses or measurements.

The "general time domain model" allows for simulation of non-linear behaviour within the contact between wheel and rail as well as within the track's superstructure. The following features may be of special interest for future application:

- the simulation of the effects of voided sleepers,
- the simulation of the effects of out-of-round wheels and
- the possibility to incorporate a non-linear material law for the ballast.

As the subgrade model is based on the approximation of receptances, the entire model yields an acceptable time exposure for the simulation despite taking into account the subgrade.

References

1. U. Gerstberger, K. Knothe, J. Siegmann and M. Yu (2000) Dynamik von Schotterfahrbahn und fester Fahrbahn – Ein Systemvergleich. Tagungsband zur Bahn-Bau '00, VDEI Verband deutscher Eisenbahn-Ingenieure, Frankfurt am Main
2. S. L. Grassie, R. W. Gregory and K. L.Johnson (1982) The dynamic response of railway track to high frequency lateral excitation. J. Mech. Eng. Sci., 24, 91–95
3. K. Knothe and M. Yu (2001) Statische und dynamische Eigenschaften von Gummi-Zwischelagen für Eisenbahnschienen. Forschung im Ingenieurwesen, 66, 247–259
4. K. Knothe and Y. Wu (1997) Erfassung der Halbraumeigenschaft des Untergrundes für die Berechnung der Fahrzeug-Fahrweg-Wechselwirkung. ILR-Mitteilung 316, Institut für Luft- und Raumfahrt, TU-Berlin
5. K. Knothe and Y. Wu (1998) Receptance Behaviour of railway track and subtrade. Arch. App. Mech., 68, 457–470
6. K. Knothe and Y. Wu (1999) Simulation des Überrollvorgangs bei diskret gelagerter Schiene. Fortschritt-Berichte VDI, Reihe 12, Nr. 412, VDI-Verlag, Düsseldorf
7. V. Koloušek (1953) Dynamik der Durchlaufträger und Rahmen. Fachbuchverlag GmbH , Leipzig
8. B. Ripke und K. Knothe (1991) Die unendlich lange Schiene auf diskreten Schwellen bei harmonischer Einzellasterregung. Fortschritt-Berichte VDI, Reihe 11, Nr. 155, VDI-Verlag, Düsseldorf
9. Y. Wu (1997) Semianalytische Gleismodelle zur Simulation der mittel- und hochfrequenten Fahrzeug/Fahweg-Dynamik. Fortschritt-Berichte VDI, Reihe 12, Nr. 325, VDI-Verlag, Düsseldorf

Measurement and Modelling of Resilient Rubber Rail-Pads

Klaus Knothe[1], Minyi Yu[2], and Heike Ilias[3]

[1] Technische Universität Berlin, Institut für Luft- und Raumfahrt,
Marchstr. 12, D-10587 Berlin
[2] Siemens Transportation Systems, Locomotive
D-80997 München
[3] Continental Aktiengesellschaft, System Engineering - Development Chassis
Systems
D-30419 Hannover

Abstract. Rail-pads are those track components which can be replaced comparatively easily in order to influence track dynamics. Quasi-static measurements as well as dynamic measurements have been performed in the low frequency range (0 - 40 Hz) and in the high frequency range (100 - 2000 Hz). The results are given as hysteresis curves and as stiffness values. In addition to secant and tangent stiffness an equivalent stiffness has been introduced. The stiffness values are frequency dependent. More significantly, they depend on the preload and on the amplitude of the harmonically varying load. Up to now a consistent time-domain model for rail-pads is not available.

1 Introduction

Introducing comments. An internal standard of DB prescribes which demands have to be fulfilled by rail-pads. To the opinion of the authors, the regulations of this standard are not sufficient. Therefore measurements have been performed in order to obtain a reliable description of the rubber rail-pad Zw 700 in the frequency range up to 2000 Hz. This rubber rail-pad is mainly used on high-speed ballasted tracks of DB. The model presented at the end is so far only valid in the frequency domain. More details of the measurements can be found in [8].

Rail fastening system and direct receptances. A typical rail fastening system W 14 together with a rail UIC 60 and a sleeper B 70 W is shown in Fig. 1. In the unloaded state the rail-pad Zw 700 has a thickness of 7 mm. The pad consists of soot-filled, foamed rubber. The pad is not profiled. Figure 2 shows a typical calculated direct receptance of the track, where ballast and subgrade as well as the pad have been modelled by spring and damper in parallel. The pad parameters mainly influence the second resonance peak.
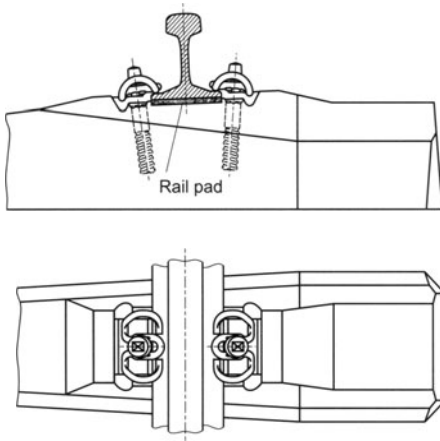
Fig. 1. Resilient rubber rail-pad Zw 700 included in Vossloh rail-fastening-system W 14 with rail UIC 60 and sleeper B 70 W

Possible measurements: Three types of measurements can be performed to investigate pad behaviour:
The first type of measurements are *field experiments* [7,6].

1. One can use the excitation of a train running on the track.
2. Or one can excite with a shaker.
3. Finally, a calibrated hammer can be used.

Field experiments seem to be the most realistic ones. However they can be extremely unreliable because of non-linearities like voided sleepers. Additionally, the values which are obtained are reliable only in the range of the second resonance peak (Fig. 2).
Laboratory measurements of pads evidently are more reliable. We decided to perform measurements of rubber pads without rail fastening system. Re-
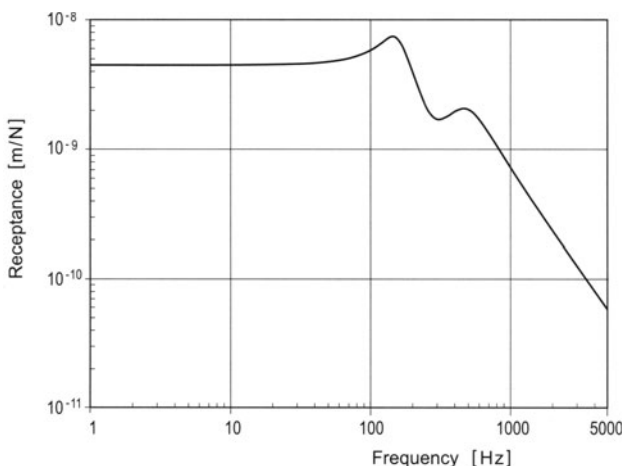
Fig. 2. Direct receptance of a continuously supported rail. Rail-pad and ballast modelled as spring and damper in parallel

sults of similar experiments are found in papers of van't Zand [10,11] and Thompson et al. [9].

The third type of measurements are *laboratory measurements of the rubber material.* Using a Finite-Element analysis the pad behaviour can be calculated for arbitrary boundary conditions. No example is known to the authors where this type of measurement and evaluation has been used for rail-pads.

2   General Concept

The typical behaviour of a rail-pad Zw 700 is shown in Fig. 3. The results of three types of laboratory measurements are compared.
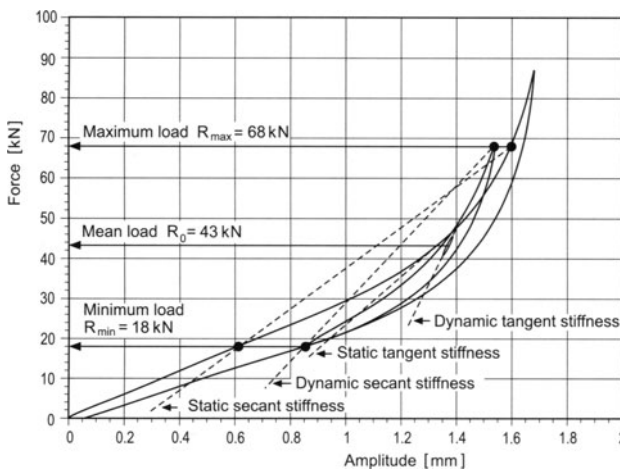
Fig. 3. Different types of loading of a rail-pad. Definition of the static secant stiffness, the static tangent stiffness, the dynamic secant stiffness and a dynamic tangent stiffness. Results of measurements for a rail-pad Zw 700 (Saargummi), 160×150×7 mm. Material of the rail-pad is foamed, soot-filled rubber

- First there is a quasi-static measurement. The rubber pad is loaded with constant loading speed between 0 and 85 kN and then unloaded. The *static secant stiffness* is defined on the loading path between 18 and 68 kN. The *static tangent stiffness* is defined at a mean load value of 43 kN on the loading path. Loading and unloading path form a banana-like hysteresis-loop.
- For low frequency experiments 43 kN are taken as a preload and the harmonical load varies between 18 and 68 kN. The *dynamic secant stiffness* which is obtained is higher than the static secant stiffness. Again loading and unloading path form a banana-like hysteresis-loop. It will be seen later that it is useful to introduce an *equivalent stiffness.*
- For high frequency laboratory experiments only force variations much smaller than 1 kN were possible. The hysteresis loop is nearly elliptical. From this ellipse a kind of *dynamic tangent stiffness* can be determined. This dynamic tangent stiffness is still higher than the dynamic secant stiffness.
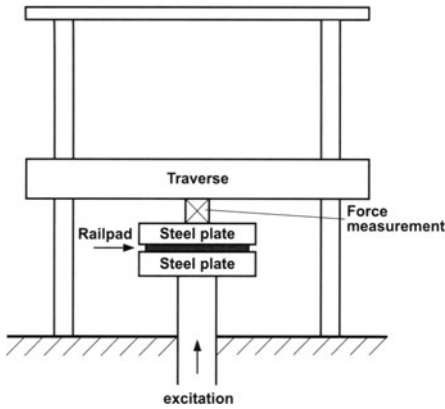
Fig. 4. Test equipment for quasi-static laboratory tests

3   Quasi-Static Experiments

The test stand which has been used for quasi-static experiments is shown in Fig. 4. The rubber pad has been excited by a hydro-pulse cylinder and the relative displacements between both steel plates have been measured for different low loading rates. The results of such quasi-static measurements are shown in Fig. 5. There are only slight deviations between the three curves. Even if both the loading and the unloading time are increased to six hours the hysteresis loop remains unchanged, indicating that the dissipation mechanism is a type of internal friction. The material behaviour is called visco-plastic [4].

It can be proved that friction between pad and steel plate does not influence the hysteresis loop qualitatively. If grease is introduced between pad and both
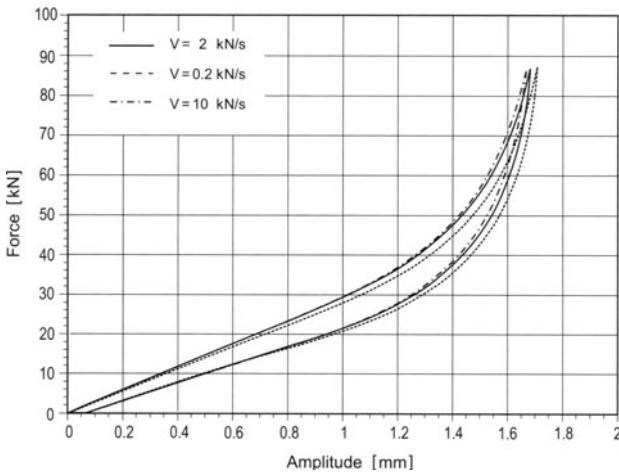
Fig. 5. Quasi-static loading of a rail-pad Zw 700 (160×150×7 mm) at three different loading rates
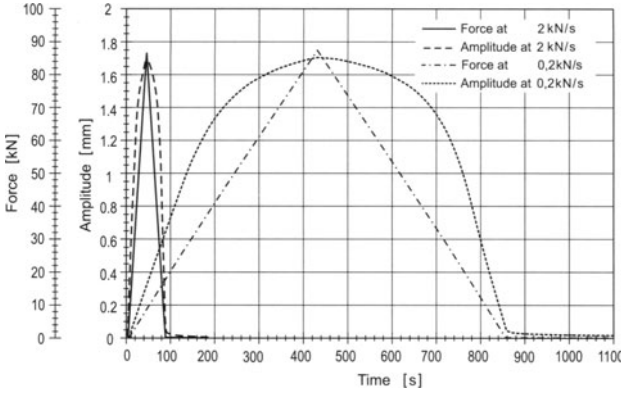
Fig. 6. Time history of load and deformation of a rail-pad Zw 700 (160×150×7 mm) during loading and unloading and after unloading

steel plates the banana-like hysteresis loop remains, however it now is much flatter and broader.

4 Low Frequency Measurements

Input variables for the measurement in the low frequency range are the frequency, the preload and the load amplitude.

In Fig. 7 and 8 hysteresis loops are shown for a preload of 43 kN and different load variation amplitudes, in Fig. 7 for quasi-static loading (0.02 Hz) and in Fig. 8 for a frequency of 10 Hz. The results are quite similar.

In order to be able to compare such hysteresis loops an *equivalent stiffness* has been introduced. First a Fourier analysis of the time history of load and displacement has been performed. The first Fourier terms have been used to determine a complex stiffness. The real part is the storage modulus and the
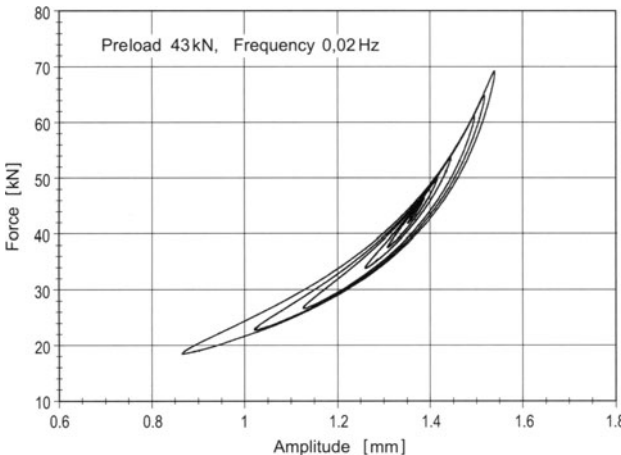
Fig. 7. Low frequency hysteresis loops for a pad Zw 700 (160×150×7 mm) with a preload of $R_0 = 43$ kN and different load amplitudes. Loading frequency $f = 0.02$ Hz
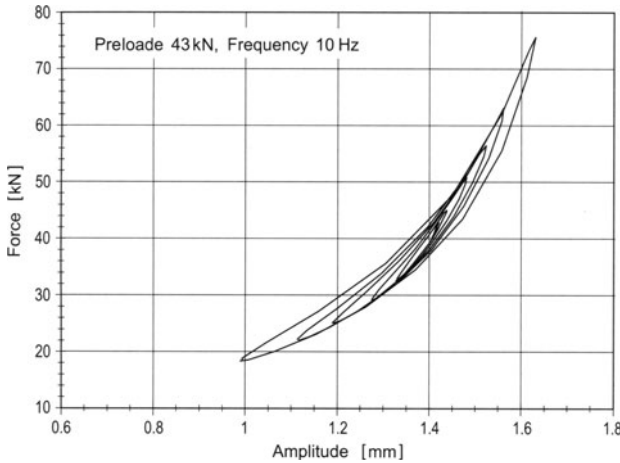
Fig. 8. Low frequency hysteresis loops for a pad Zw 700 ($160 \times 150 \times 7$ mm) with a preload of $R_0 = 43$ kN and different load amplitudes. Loading frequency $f = 10$ Hz

imaginary part is the loss modulus. The amplitude of both is called *equivalent dynamic stiffness*. If only the first Fourier terms are used, then instead of the banana-like hysteresis loops, elliptical hystereses are obtained (Fig 9).

Results of low frequency measurements are presented for a preload of 43 kN in Fig. 10. For a load amplitude of 25 kN the dynamic secant stiffness and the equivalent dynamic stiffness are nearly the same. For very small load amplitudes the equivalent dynamic stiffness is higher by a factor of more than 2 and equals the dynamic tangent stiffness!
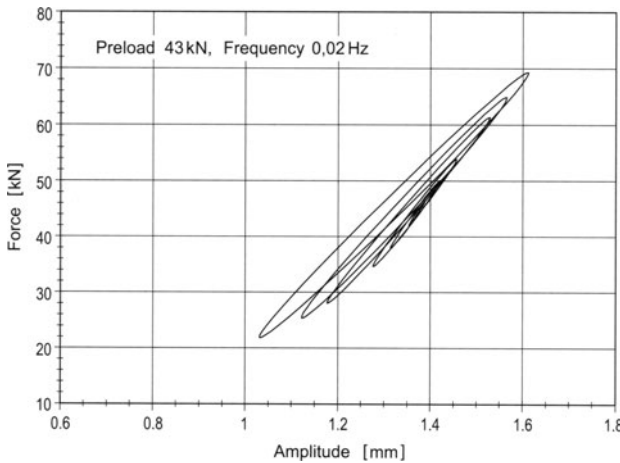
Fig. 9. Equivalent ellipses for the banana-like hysteresis loops of Fig. 7
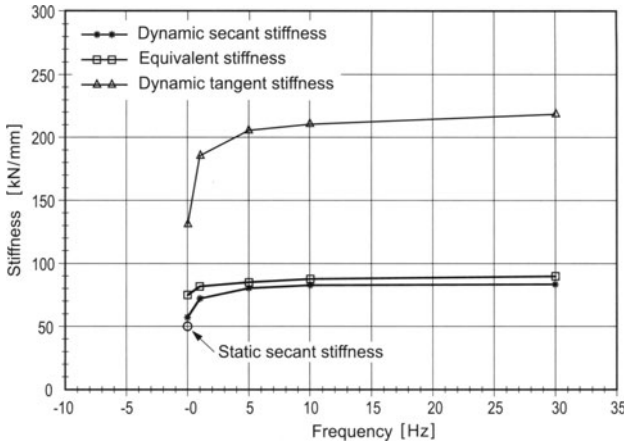
Fig. 10. Comparison of different stiffness values for a rail-pad Zw 700 (160×150×7 mm) under a preload $R_0 = 43\,\text{kN}$

5 High Frequency Experiments

For high frequency measurements different equipment has been used, c.f. Fig. 11. Without going into detail it should be mentioned that there were two pads placed between three steel-plates. Details can be found in [8].

The complex stiffness for frequencies between 1 and 2000 Hz are shown in Fig. 12. On the horizontal axis a logarithmic scale is used for the frequency. On the upper part the storage modulus is given and on the lower part of the figure the loss modulus is shown. In addition to the measured values, fitted lines have been determined. The increase of both moduli with increased frequency is obvious. The increase is much more dominant for the loss modulus than for the storage modulus which indicates that visco-elastic effects are more pronounced for higher frequencies.
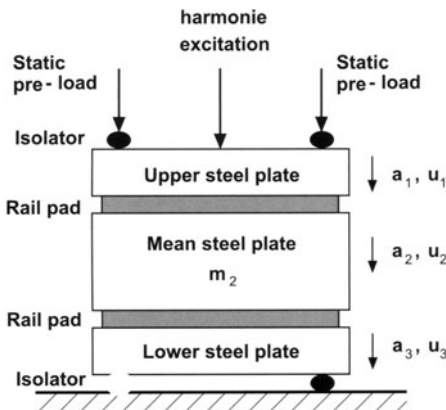
Fig. 11. Arrangement of the test equipment for experiments in the medium and high frequency range
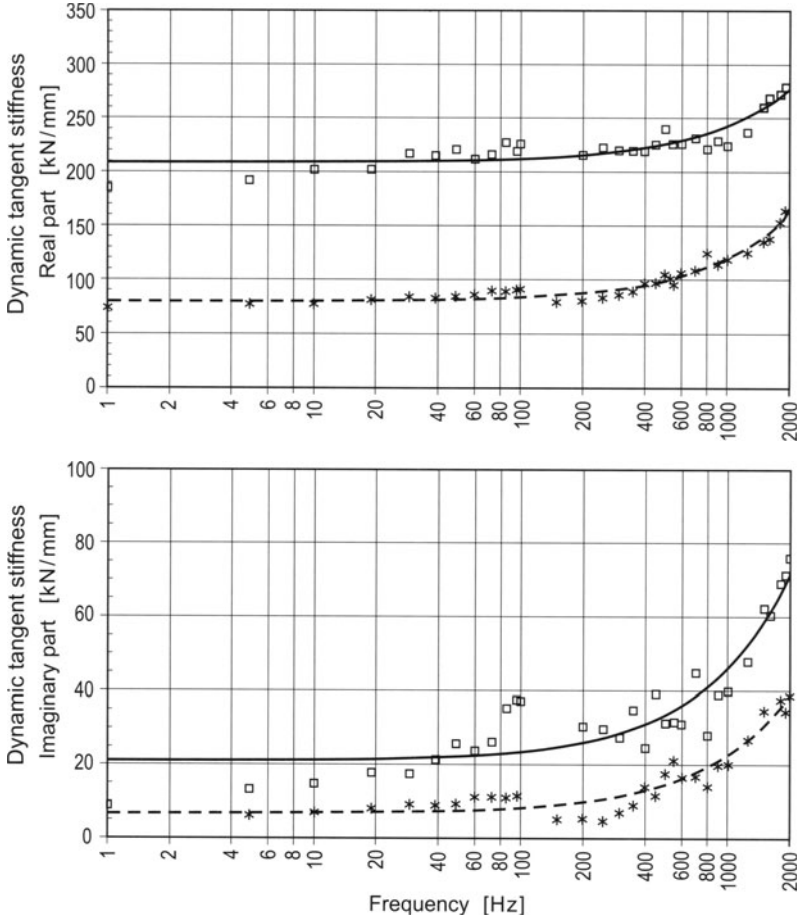
Fig. 12. Dynamic tangent stiffness of the rail-pad Zw 700 A60 SGW 95 (dimensions 160×150×7 mm). Summary of results of low and high frequency measurements with preload R_0 of 25 kN (∗) and 43 kN (□) respectively. According to measurements of M. Yu (Sept. 2000) with corresponding fitted lines

6   Interpretation of Measurement Results

The results of the of the measurements can be summarized and interpreted as follows

- The quasi-static loading path can be described by hyper-elastic material behaviour, e.g. by an Ogden model.
- For low frequency cyclic loading the rubber material shows visco-plastic behaviour.
- Additionally, for higher frequencies viscoelastic effects are important.
- The equivalent complex stiffness depends on the frequency, on the preload and on the amplitude of the periodically varying load!

- The equivalent stiffness increases with increasing preload and decreasing amplitude.

7 Frequency Domain Model of Rail-Pads

Up to now modelling is only possible in the frequency domain. As already mentioned, the equivalent complex stiffness depends on the frequency f, on the preload R_0 and on the amplitude ΔR of the varying load. This complex function has to be approximated.

- It is assumed that for each frequency the equivalent complex stiffness can be approximated by a bi-linear function with respect to preload and load amplitude. This should be done for two frequencies, namely 30 and 1000 Hz.
- Additionally, it is assumed that the parameters depend linearly on the frequency. This linear dependency on the frequency is justified as it was possible to determine fitted lines.

For such a simple approximation at least eight values are necessary for both the storage and the loss modulus. The data for the storage modulus is shown in Table 1. Two of these values had to be estimated very roughly as up to now laboratory measurements with high amplitudes are not possible in the high frequency range. As soon as more measurements are available, the model can be refined.

Table 1. Data for the tri-linear approximation of the storage modulus $k'_\mathrm{p} = k'_\mathrm{p}(f, R_0, \Delta R)$ of a rail-pad Zw 700 according to measurements of M. Yu

Amplitude	Frequency $f = 30$ Hz		Frequency $f = 1000$ Hz	
	$R_0 = 25$ kN	$R_0 = 43$ kN	$R_0 = 25$ kN	$R_0 = 43$ kN
	kN/mm	kN/mm	kN/mm	kN/mm
$\Delta R = 25$ kN	68*	91	(103)**	(123)**
$\Delta R \ll 1$ kN	85	210	123	240

* Estimated from secant stiffness.
** Rough Extrapolation. No measurements available.

8 Concluding Remarks

Finally some comments are necessary with respect to the time domain model. What is first needed here is a constitutive law. Here, the work of Haupt [3–5] has to be mentioned. Another possibility is the model of Gudehus which

has been developed to model ballast behaviour [2,1]. The results from measurements which have been made for ballast material are qualitatively quite similar to the results which have been obtained for pads. The reason may be that on the microscopic level the pad material is not at all homogeneous due to the filling with microscopic soot particles and foam bubbles.

References

1. Gudehus. G. (1997) Attractors, percolation thresholds and phase limits of granular soils. Proc. Powders and Grains, 1–15
2. Gudehus. G. (1997) A comprehensive constitutive equation for granular material. Soils and Foundation 36:1–12
3. Haupt. P. (1995) On the thermodynamic representation of viscoplastic material behavior. Mitteilung des Instituts für Mechanik 3/1995, Universität Kassel – Gesamthochschule
4. Haupt. P. (2000) Continuum Mechanics and Theory of Materials. Springer, Berlin
5. Haupt P., Sedlan K. (2001) Viscoplasticity of elastomeric materials: experimental facts and constitutive modelling. Archive of Applied Mechanics 71:209–262
6. Knothe. K. (2001) Gleisdynamik. Ernst & Sohn, Berlin
7. Knothe K., Grassie S.L. (1993) Modelling of railway track and vehicle/track interaction at high frequencies. Vehicle System Dynamics 22(3–4):209–262
8. Knothe K., Yu M. (2001) Statische und dynamische Eigenschaften von Gummi-Zwischenlagen für Eisenbahnschienen. Forschung im Ingenieurwesen 66:247–259
9. Thompson D. J. , Nijman E.J.M. (1991) Railway rolling noise. Measurements of the dynamic properties of five rail fastener systems. Report for the C163 ORE Committee, TNO
10. Van't Zand J. (1990) Experimental evaluation of dynamic properties of rail-pads at lower temperatures (revised copy). Report No. 7-88-105-8, Delft University of Technology, Department of Civil Engineering, Road and Railroad Research Laboratory
11. Van't Zand J. (1994) Assessment of dynamic characteristics of rail-pads. Rail Engineering International Edition, 15–17

Model-Based Investigation of the Dynamic Behaviour of Railway Ballast

Holger Kruse and Karl Popp

University of Hannover, Institute of Mechanics,
Appelstr. 11, D-30167 Hannover, Germany

Abstract. The representation of ballast in conventional track models is unsatisfactory, since still important physical phenomena inside the ballast layer are not understood. Therefore, in this paper an alternative theoretical approach, the Molecular Dynamics method (MD method) as a powerful tool for the investigation of granular material, is investigated. In contrast to a continuum description, in the corresponding 2D-model each single stone of the ballast can be taken into account. In the frame of a numerical time-step integration, the contact forces are calculated from the overlap area of the particle geometries.

The paper includes information about the sensitivity of the model behaviour on initial conditions and contact law parameters. Furthermore, the contact network, the quasi-static stiffness of the ballast layer and its long-term behaviour are addressed. Particular emphasis is put on the description of current difficulties and challenges in applying the MD method.

1 Introduction

Due to its good draining properties, ballast has played an important role already in pioneering days of railway technology. However, even today the dynamic processes inside the ballast layer are not understood in detail. Since during the last years railway companies have faced enormous maintenance costs caused by ballast degradation especially on high-speed lines [9], a lot of research has been initiated to get a better insight into the corresponding damage phenomena.

The original objective of the research project Po 136/20 — financially supported by the 'Deutsche Forschungsgemeinschaft' (DFG) — was the development of improved track models based on wave solutions. In this context, the software package TTI has been originated. After giving an overview on the consideration of the ballast layer in conventional track models in chapter 2, chapter 3 provides some information about the capacity of TTI. Furthermore, a few results are listed. The main focus of this paper, however, is put on chapters 4–6 which deal with the Molecular Dynamics method.

The MD method has first been applied to railway ballast by Estradé in 1989 [11]. He investigated a system similar to that depicted in Fig. 1. As the computational power was quite low before 1990 compared to today, only a circular shape for the ballast stones could be used and a few simulations

could be performed. They already gave some insight into the dependence of the contact forces between the ballast stones on the coefficient of friction or the pressure distribution below the sleeper.
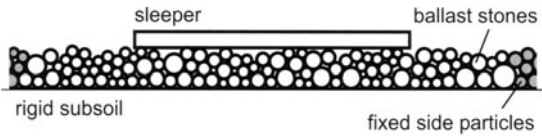
Fig. 1. Model of ballast layer below sleeper, from [11]

Later on, some work has been done in France as part of the European research programme EUROBALT, but without substantial outcome. Hence, the results presented in chapter 6 should be understood as further steps to assess, whether the MD method can become a powerful tool for engineers in optimizing the design of ballasted tracks.

2   Railway Ballast in Conventional Track Models

Up to the late eighties, most scientists modelled ballast and subsoil together by a viscoelastic foundation below the sleepers, see e.g. [12] and Fig. 1a. If the wave propagation in the soil was analysed, first analytical or semi-analytical approaches were made, which were based on a halfspace representation of ballast and subsoil (Fig. 1b). Besides, in the beginning of the nineties, a few authors applied discrete mass-spring-damper models for the ballast, similar to that depicted in Fig. 1c.
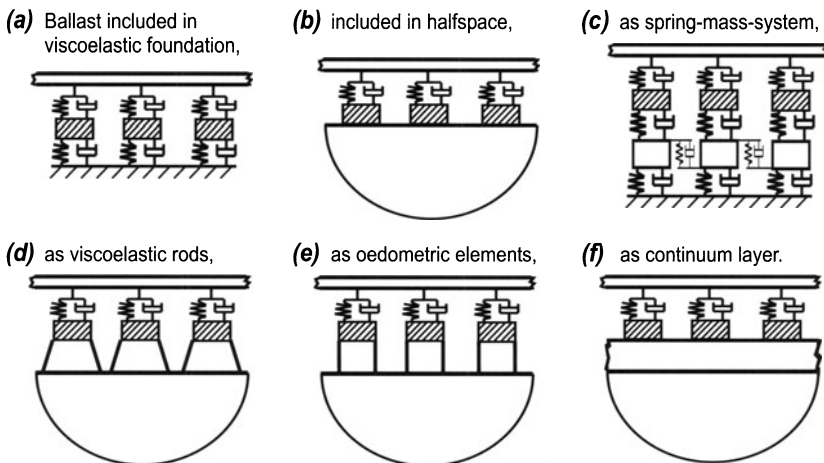
Fig. 2. Different representations of ballast and subsoil in existing track models

During the last five years, further types of ballast models have been introduced: the viscoelastic rod model (Fig. 1d) by Knothe [4], the oedometric element model (Fig. 1e) by Gudehus et al. [14] and the continuum model (Fig. 1f) by Savidis et al. [1].

However, even today it is unknown which of these models gives the best agreement with the real structure. Still many open questions exist regarding a correct formulation of the boundary conditions between sleeper and ballast or ballast and subsoil, regarding a possible wave propagation inside the ballast layer or regarding a proper consideration of damping. And the long-term behaviour of the ballast is even less understood.

3   The Software Package Train Track Interaction (TTI)

TTI is a software tool for the investigation of both, single track components and full train-track systems. A typical track model is shown in Fig. 3. It fulfills the two basic requirements of linearity and periodicity with respect to the track's length direction. Hence, the rails and the subsoil have infinite extension. For the rails the Timoshenko beam theory is used. The pads are taken into account by massless elements which can have an arbitrary frequency dependent complex stiffness.
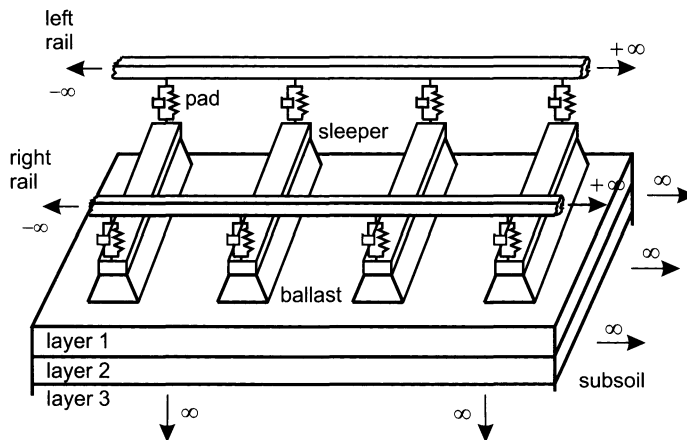
Fig. 3. Typical track model used in software TTI

A consideration of the elasticity of the equally spaced sleepers is possible. As already depicted in Fig. 1d, below the sleepers there are located viscoelastic rods representing the ballast. Therefore, the ballast parameters like geometric dimensions, stiffness or density are physically motivated. The subsoil model can be of arbitrary complexity. It just has to be linear like all other components. In Fig. 3 a layered halfspace is shown, which leads to a sleeper coupling via the substructure.

Concerning the vehicle models, simple systems like a single point mass or an oscillator as well as advanced multibody systems with elastic components, e.g. deformable wheelsets, can be coupled to arbitrary track models. The connection to the track is realised by a linear Hertzian spring in vertical direction. As source of excitation, TTI provides the discrete rail support itself, unbalances in the wheelsets, unround wheels or periodic rail-surface roughnesses.

Since the whole train-track system is described by linear partial or ordinary differential equations, the vertical dynamics in the steady state can be calculated using transformations in frequency and wavenumber domain. The modular concept in defining the model structure is additionally implemented in a complex network of algorithms. The most intricate algorithm, whose fundamental steps are sketched in Fig. 4, is needed for the determination of the track motion under moving forces harmonically varying in time. For a detailed explanation we refer to the literature [5–8], where also information about other parts of the algorithm can be found.
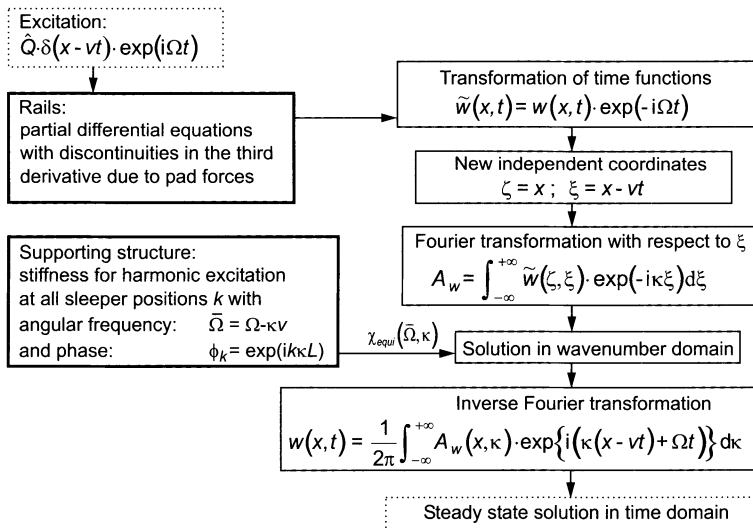
Fig. 4. Basic steps of the algorithm for a periodic track loaded by moving forces

With TTI, many parameter studies have already been performed to investigate the influence of the discrete rail support, of the sleeper elasticity, the shear deformation of the rails and of different excitation sources. Further research was addressed to rigid track models, the importance of a correct description of the dynamic pad properties or the consequences of taking into account the wheelset elasticity, see again [5–8].

Since this paper focuses on ballast and subsoil, in the following a few results regarding these components are summarised. Here, emphasis is put on the mid-frequency range.

- The track receptance and the dynamic wheel/rail contact force depend strongly on the chosen ballast and subsoil model. Comparing the model in Fig. 3 to a track model with Winkler foundation, even the qualitative shape of the corresponding frequency response functions is changed.
- In the frequency range below 350 Hz, the system reacts very sensitively especially to a variation of subsoil parameters. Different layered halfspaces can lead to totally different locations of the resonances and deviations in the curves up to 100 percent.
- The sleeper coupling via the subsoil changes the shape of the rail bending line. The new results better fit to experimental findings.
- Special resonance effects occur for very high vehicle velocities close to the velocity of Rayleigh waves in the subsoil. Here, unexpectedly strong vibrations can occur.

4   Molecular Dynamics Method

The MD method is based on scientific works which appeared in the years before 1980. One of the pioneers was Cundall [2]. In contrast to the common multibody dynamics approach, he did not focus on a single mathematical model for the whole multi-particle system in the form of a system of differential equations due to the following reasons:

- For the multi-particle systems investigated, a frequent formation of new contacts takes place and existing contacts open. In multibody dynamics, both events require an update of the governing equations.
- Even when the contact network remains the same, in each contact a transition between sticking and slipping is possible, which also leads to a change in the structure of the system of equations of motion.

These difficulties can be avoided if the principles of momentum and moment of momentum are formulated separately for each particle, a procedure, which is possible if nearly no kinematic constraints between different particles exist. Hence, most particles keep all their translational and rotational degrees of freedom. The penetration of one particle into another one is avoided by introducing very high stiffnesses which simultaneously allow for a rough consideration of particle deformations. Performing a time-step integration, the system dynamics are determined.

Starting with the system state at time t, which is defined by position and velocity of all particles, the situation at time $t + \Delta t$ is calculated as follows:

1. Detection of all contacts using an adequate search algorithm.

2. Calculation of contact forces and assignment to the corresponding contact partners.
3. Determination of the accelerations of all particles by using the principles of momentum and moment of momentum.

This methodology allows for the investigation of very large systems with hundreds of thousands of particles. Some typical examples are shown in Fig. 4 [13]. The different implementations of the MD method do not only differ in the search algorithms and integration routines used, but also in the shape of the particles and the applied contact laws.
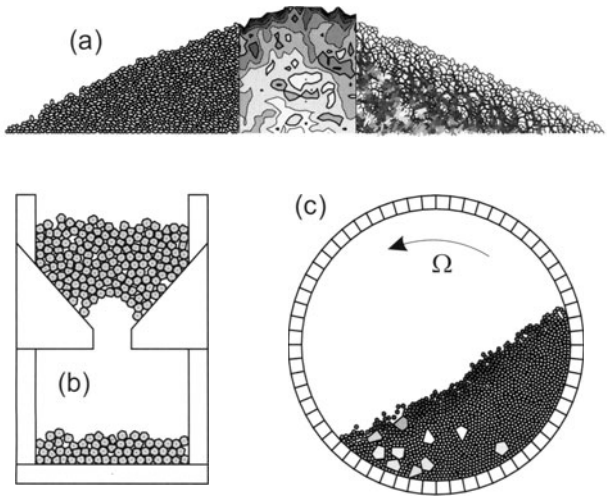
Fig. 5. Applications of the MD method. (**a**) Particle configuration (left), pressure (middle) and contact force network (right) in a sand heap. (**b**) Choked funnel. (**c**) Particle motion in a mixing drum

The main disadvantage of the MD method is the required very small constant time-step Δt due to the high stiffnesses mentioned above. The time-step Δt has to be chosen in a way that both impacts and stick-slip transitions can be treated like continuous processes. Considering the time-consuming operations which have to be performed during each integration step, enormous computational effort is required.

5   Application to Railway Ballast

Many granular materials consist of very small grains. Thus, an upscaling of the geometry is necessary which often leads to difficulties related to the interaction with boundaries or an adaptation of forces or material parameters. In the case of railway ballast such an upscaling is not needed. Here, the main

challenge is to get an adequate representation of the real stone shape, which affects the very small particle migrations causing settlement. The use of round particles seems to be insufficient.

In the following subsection the model geometry is described in detail. Subsequently, further subsections give insight into the contact laws, the simulation sequence and input and output data, respectively. A description of the algorithm for the contact detection and the algorithm for the determination of the overlap areas can be found in [10,8]. As integration routine the Gear predictor corrector formula of fifth order is used [8].

5.1   Model Geometry

The 2D-model is depicted in Fig. 2. In these first studies, it does not take into account a complete ballasted track. Due to the fixed wall particles, it rather describes the conditions during so-called ballast box tests which have been performed by Selig [15] and others. The dimensions of the ballast layer are chosen in such a way that the model could also be interpreted as a side view of a single sleeper resting on ballast. However, then the boundary should provide similar dynamic properties like the subsoil in reality, which is possible in the frame of a MD simulation.
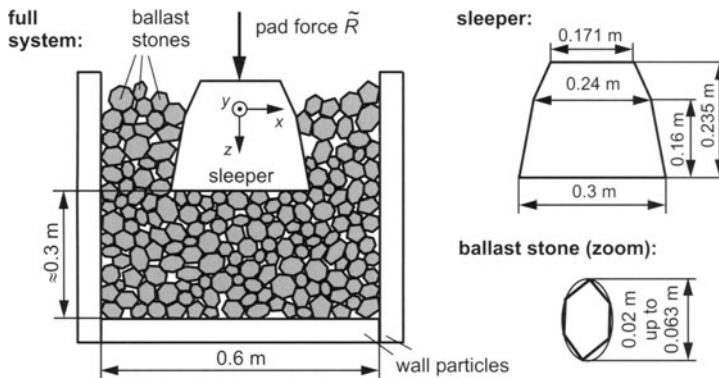
Fig. 6. Model of single sleeper resting on ballast

The ballast stones are represented by polygons with a variable number n of corners, $n = 5 \ldots 8$. These polygons are inscribed into ellipses with semiaxes $l_{a/b} = 0.01\,\mathrm{m} \ldots 0.032\,\mathrm{m}$. The shape of the sleeper corresponds to the mean cross section of a common German B70-sleeper. While the ballast stones keep all their three degrees of freedom, the sleeper can only move vertically.

5.2   Contact Forces

The contact forces are determined by evaluating the contact geometry which is sketched in Fig. 3. It is evident that important features of the shape of

real ballast stones — like their roughness or concave surface sections allowing for multi-point contacts — are not taken into account. Furthermore, even for simple geometries the knowledge about the contact laws regarding lateral stiffness, micro-slip and damping is poor. Hence, a quite high uncertainty with respect to the contact force calculation has to be accepted. Therefore, it must be proved carefully, whether the model behaviour is sufficiently robust concerning a change in the contact laws. The results of corresponding investigations are given later.
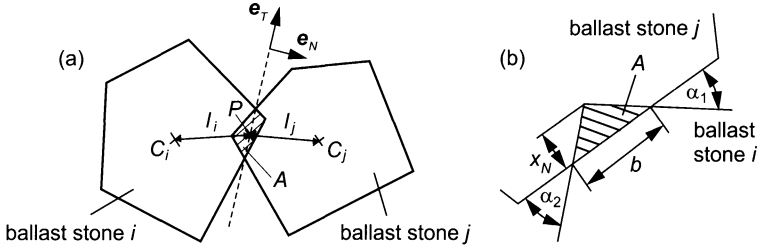
Fig. 7. Contact geometry

For each contact, we choose a force vector which acts in point P and can be divided into four components regarding stiffness and damping,

$$\boldsymbol{F} = (F_{N,C} + F_{N,D}) \cdot \boldsymbol{e}_N + (F_{T,C} + F_{T,D}) \cdot \boldsymbol{e}_T. \tag{1}$$

The directions of the unit vectors $\boldsymbol{e}_N$ and $\boldsymbol{e}_T$ are shown in Fig. 3a. The spring force in normal direction reads

$$F_{N,C} = \frac{\tilde{E}A}{l} \approx \frac{b}{2l} \cdot \tilde{E} \cdot x_N, \tag{2}$$

where $A \approx 0.5 \cdot b\,x_N$ is the overlap area of the undeformed polygons with b as contact length and x_N as penetration depth (see Fig. 3b), $\tilde{E}$ is Young's modulus times the system depth of 1 m and $l = 4 \cdot l_i \cdot l_j \cdot (l_i + l_j)^{-1}$ is a measure for the size of the stones in contact. Hence, with decreasing particle size the contact becomes stiffer — a consequence of the reduction of deformable material. Furthermore, it is worth to be mentioned, that in nearly all cases Eq. (2) results in a quadratic relation between force and particle approach. This nonlinear stiffness is used to take into account the high surface roughness and elastic properties of the real 3D-system. The damping force is calculated by

$$F_{N,D} = \begin{cases} 2D_N \cdot \sqrt{\tilde{E}m_N/bl} \cdot \mathrm{d}A/\mathrm{d}t & \text{for } \mathrm{d}A/\mathrm{d}t > 0 \\ \max\{2D_N \cdot \sqrt{\tilde{E}m_N/bl} \cdot \mathrm{d}A/\mathrm{d}t, -F_{N,C}\} & \text{for } \mathrm{d}A/\mathrm{d}t < 0 \end{cases}, \tag{3}$$

where $m_N = m_i \cdot m_j \cdot (m_i + m_j)^{-1}$ depends on the masses of the two contact partners and D_N denotes the damping ratio. The upper limit of $F_{N,D}$ during

a separation ($dA/dt < 0$) ensures that no attractive forces act. By Eqs. (2) and (3), it is achieved that any small vibrations in normal direction around an arbitrary prestressed state are damped by the same damping ratio D_N. This damping includes material damping due to stone deformations and the energy loss due to sound radiation and micro-slip in the contact zone.

The tangential force is directed to reduce the tangential relative motion between the particles. Due to the Coulomb's law, an upper limit exists,

$$|\boldsymbol{F}_T| = \min(\mu \cdot |\boldsymbol{F}_N|, F_{T,C} + F_{T,D}), \tag{4}$$

where μ denotes the coefficient of friction. No distinction between static and dynamic friction is made. Sticking is taken into account by incrementally adapting the spring force

$$F_{T,C}(t) = F_{T,C}(t - \Delta t) + \tilde{G} \cdot b/l \cdot v_T \cdot \Delta t, \tag{5}$$

where v_T is the relative tangential velocity. The factor b/l introduces a coupling between the time variant normal and tangential stiffnesses. In the case of a constant normal stiffness ($b/l =$ const.), the tangential spring is linear. The damping component reads

$$F_{T,D} = 2D_T \cdot \sqrt{\tilde{G} \cdot b/l \cdot m_T} \cdot v_t \tag{6}$$

with $m_T = (1/m_i + 1/m_j + l_i^2/I_i + l_j^2/I_j)^{-1}$, which also takes into account the moments of inertia I_i and I_j. Again, the damping ratio D_T is valid for small vibrations around an arbitrary contact situation. The nominal parameters for the contact laws are given in Fig. 6. The value of D_N has been estimated from drop tests and is also used for D_T, $D_N = D_T = 0.5$. The values for the stones' Young's modulus and density correspond to the actual material basalt. Their units are adapted due to the fact that a 2D-model is used. As relation between $\tilde{G}$ and $\tilde{E}$ the number '2/5' is chosen.

5.3   Input and Output Data

As input, the vertical force $\tilde{R}$ acting on the sleeper has to be specified. In the case of simulating a laboratory test this force is usually known. If a train passing is investigated, it can be determined from either measured or calculated pad forces. Since, for a real track, the dynamic pad forces differ strongly between the individual sleepers, the use of calculated pad forces leading to mean values is preferable. The time function of $\tilde{R}$ together with the vehicle model applied is depicted in Fig. 7. The force $\tilde{R}$ has been determined with the software TTI, which is described in chapter 3.

As output, sleeper force, sleeper displacement, mean velocity of the ballast stones, their migration, kinetic and potential energies and some statistic information about the contact forces are recorded during the program runs. In addition, the system configuration is stored at equally spaced time points so that animations may be generated.

● Approximation of shape

no roughness

no concave shape;
no multi-point contact

$\tilde{\rho} = 5000$ kg/m²

● Approximation of damping

$$D_N = D_T = 0.5$$

drop test

first contact with ground

contact length b

stone size l

● Approximation of friction

$$\mu = \mu_0 = 0.3$$

● Approximation of normal and tangential contact stiffness

$$c_N = b/l \times \tilde{E} = b/l \times 6 \times 10^{10} \text{ N/m}; \quad c_T = b/l \times \tilde{G} = 2/5 \times c_N$$
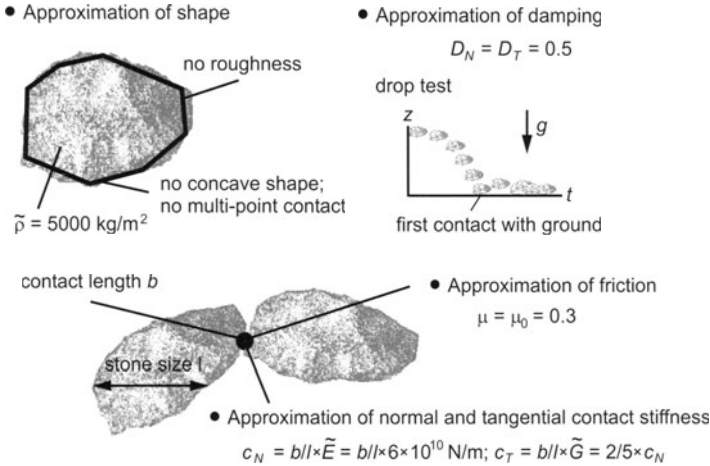
Fig. 8. Contact law — characteristics and parameters

axle load:
32 t

$v = 100$ km/h

6 m   6 m
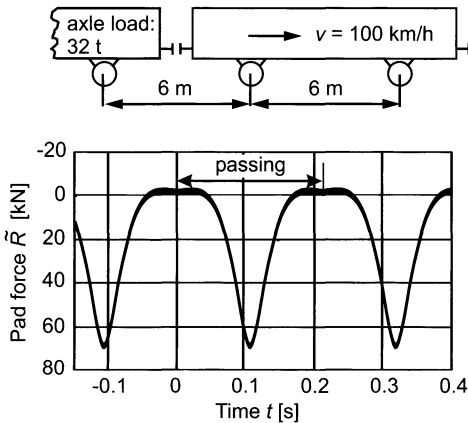
Fig. 9. Time dependent pad force and vehicle model used

5.4 Simulation Steps

For each simulation, a sequence of steps has to be carried out to arrive at a system state that is comparable to that in reality. These steps are sketched in Fig. 8.

1. Specification of ballast stone shape by a random number generator. Positioning of the stones without overlaps.
2. The ballast stones fall down due to an initial vertical velocity and due to gravity. In the ballast layer a certain contact network arises.
3. Smoothing of the surface by removing some of the top stones.
4. The sleeper and further stones at both sides of the sleeper are added in an identical procedure, see steps 1 and 2.
5. Optionally, a tampering process can be simulated.
6. The load cycles are applied.
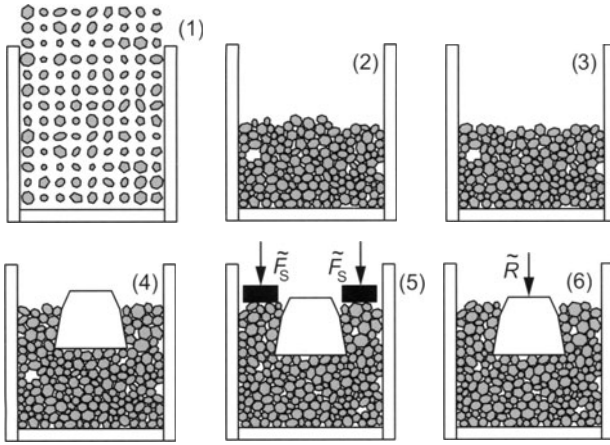
Fig. 10. Simulation steps

The required time-step for the integration is influenced by the mass m_{min} of the smallest stone and Young's modulus $\tilde{E}$. The equation

$$\Delta t < 0.3 \cdot \sqrt{m_{min}/\tilde{E}} \tag{7}$$

yields a rough upper limit which is fulfilled by $\Delta t < 1\mu$s. However, it has to be checked carefully whether this time-step guarantees a sufficiently accurate discretisation of all vibrations including tangential ones and rotational motions. The results presented later have been obtained with $\Delta t = 0.5\mu$s.

6   Simulation Results

To examine whether the MD method can provide reliable insight into the ballast behaviour, first the plausibility of the simulation results and their robustness regarding the initial and boundary conditions and the contact law parameters have to be investigated. This is done in chapter 6.1. Subsequently, chapter 6.2 is addressed to the short-time dynamics of the ballast, before chapter 6.3 gives some idea of an analysis of the long-term behaviour including abrasion and stone migration. A more detailed study of these three topics can be found in [8].

6.1   Plausibility and Robustness Tests

Kinetic and Potential Energy Fig. 9 shows the kinetic (E_{kin}), gravitational (E_{pot}) and contact energy (E_{con}) during a whole simulation run, all three of them summarised over all ballast stones. The single simulation steps sketched in Fig. 8 can be identified clearly. In step 2 ($t < 0.7$ s) E_{kin} rises first, starting already from a quite high value because of the initial stone velocities. Approaching $t = 0.7$ s more and more stones reach the ground and

come to rest. Hence, E_{kin} drops and the steady decrease of E_{pot} ends. For $t = 0.7\,$s both energy forms jump due to the addition of further particles (step 4) and an analogous process takes place. During the wheelset passings (step 6) E_{kin} and E_{pot} nearly stay constant, since the stone motion is now extremely small. While E_{con} was negligible before, the dynamic load causes the highest fluctuations here. The time dependence — see detailed view of Fig. 9 — is very similar to that of $\tilde{R}$, cp. Fig. 7. The smooth curve shape also means that the model behaviour is not disturbed by numerical instabilities.
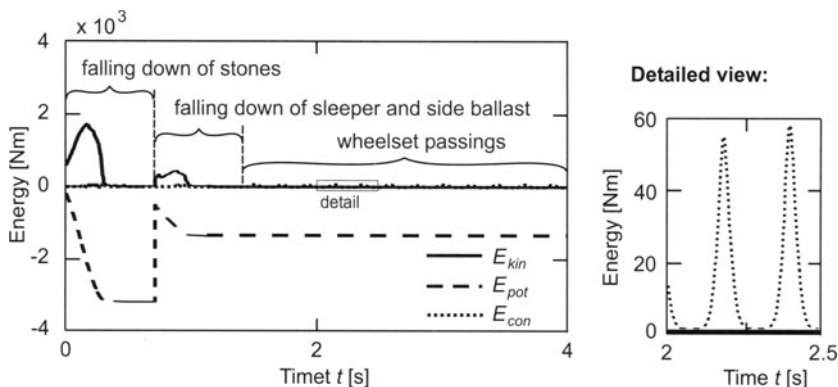
Fig. 11. Kinetic, gravitational and contact energy as functions of time

Influence of Initial Conditions Since the ballast stone shape is determined by a random number generator, even with identical system parameters, the individual simulations differ significantly, see Fig. 10. Besides, because of the stone corners, the falling down of the ballast is a chaotic process, which depends sensitively on the initial state. Thus, a prediction of the formation of cavities is impossible.

For each of the diagrams in Fig. 10 ten calculations have been performed. In diagram (a) no tampering was simulated, in (b) the ballast layer was slightly and in (c) highly precompressed. Again, during the load cycles the pad force depicted in Fig. 7 was applied.

The reversible and the irreversible part of the vertical sleeper motion during the single wheelset passings can be seen clearly. Only the latter contributes to the settlement. During the first few load cycles, this settlement is quite high and decreases rapidly. The big differences between the curves are obvious. Looking at diagram (a), after 20 load cycles the sleeper has settled by 11 mm in one simulation, in another simulation just by 2 mm. The following conclusions can be drawn:
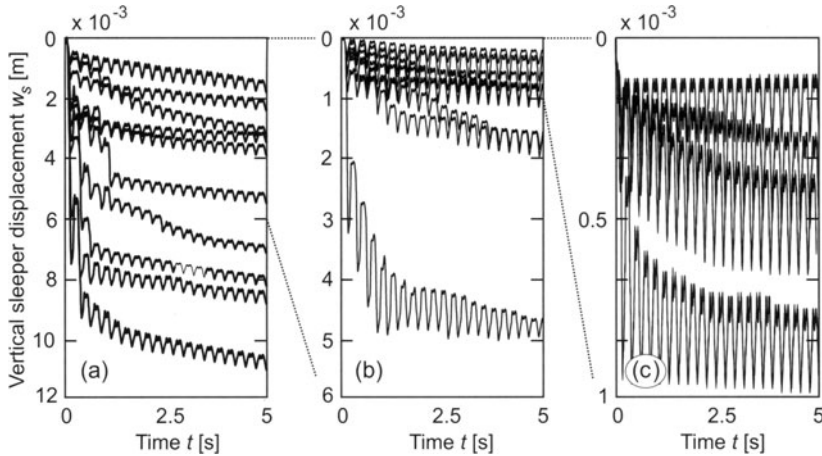
Fig. 12. Sensitivity of ballast settlement during the first load cycles due to varying initial conditions. (**a**) Not precompressed. (**b**) Slightly precompressed. (**c**) Highly precompressed

- The model shows a behaviour, which can also be observed for a real ballast layer and which probably plays an important role in discontinuous ballast settlement.
- The characteristics of a system with one set of parameters cannot be obtained on the basis of a single simulation. At least 40 up to 100 program runs should be performed allowing for the determination of reliable mean values and standard deviations.
- The high relative deviations in the settlement behaviour cannot be reduced by tampering, though the mean settlement decreases strongly.

Influence of Boundary Conditions In extended parameter studies it becomes obvious that most changes in the boundary conditions — like e.g. a variation of the stiffness and the friction coefficient of the wall particles or a variation of the model dimensions — do not affect the sleeper displacement noticeably [8]. Surprisingly, this was even observed concerning a transition in the stone shape from almost round to flattened and cornered.

A very interesting study is the investigation of the configuration depicted in Fig. 11. Here, the length view of the sleeper is modelled. After a compression phase with slow load cycles ($t < 8\,\mathrm{s}$) some particles on the right side are removed. The bold line in Fig. 11 and the right sketch of the model constellation ($t = 15\,\mathrm{s}$) make clear that the ballast layer remains stable without any side wall preventing the stones from sliding down.

Influence of Contact Parameters In Fig. 12, each curve represents the vertical sleeper displacement averaged over ten simulations. Starting with the
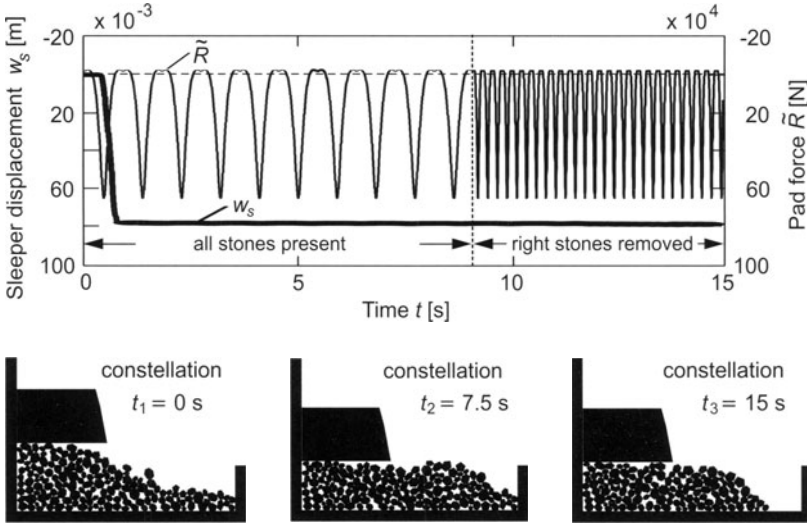
Fig. 13. Length view of sleeper — system behaviour and model constellation at different time points

nominal system — the corresponding contact parameters are given in Fig. 6 — the two stiffnesses c_N and c_T and the two damping ratios D_N and D_T are separately reduced by a factor of five. After 20 wheelset passings, even for such a strong change in the contact parameters the irreversible sleeper displacement differs by less than 25 %. Especially the influence of damping seems to be low. A reduction of c_T leads to a slight decrease in the ballast stone migrations, since stick-slip transitions occur less frequently. Besides, the reversible sleeper elongation increases which is even more pronounced for $c_N = 0.2 \cdot c_{N,nominal}$.
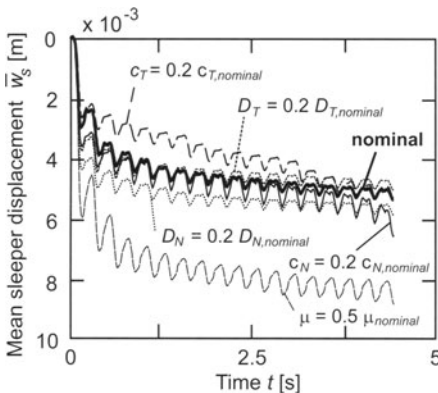
Fig. 14. Vertical sleeper displacement for different contact law parameters

Figure 12 also contains a curve corresponding to a simulation where the friction coefficient μ has been reduced by a factor of two. It is obvious that this contact parameter plays a more important role. The reduction yields nearly twice the settlement that occurs in the nominal system. In reality, the friction value can change due to moisture.

6.2   Short-Time Dynamics

Contact Force Network A typical contact force network is depicted in Fig. 13. The width of the lines connecting the centers of mass of the particles in contact are proportional to the normal contact force. The small lines inside represent the amplitude of the tangential contact forces. It is clearly seen that the pad force is distributed among only a few load paths. Hence, the load of the individual ballast stones differs significantly. In the 3D-case a more homogeneous pressure distribution may occur.
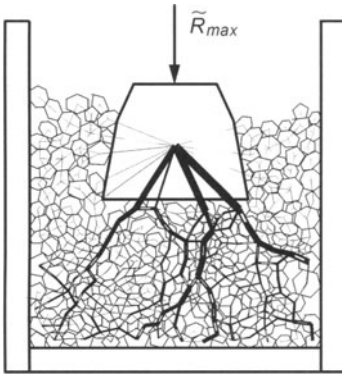
Fig. 15. Contact force network

Quasi-Static Stiffness of Ballast Layer The influence of inertia and viscous damping forces can be neglected if a pad force is chosen like depicted in Fig. 14a. Here a load cycle takes 2 seconds, which is about ten times longer than the duration of the load cycles before. The resulting sleeper displacement is given in Fig. 14b. As the load is applied during the initial compaction phase of the ballast layer, still some irreversible sleeper deflection can be seen.

Fig. 14c shows the pad force versus the sleeper displacement. The hysteresis curve allows for a determination of the energy loss per load cycle,

$$E_V = \overset{loading(1,3)}{\int} \tilde{R}(w_s)\,\mathrm{d}w_s + \overset{unloading(2,4)}{\int} \tilde{R}(w_s)\,\mathrm{d}w_s\,. \tag{8}$$

Since viscous damping can be neglected, the energy loss is due to friction between the stones. Finally, the quasi-static stiffness of the ballast layer as
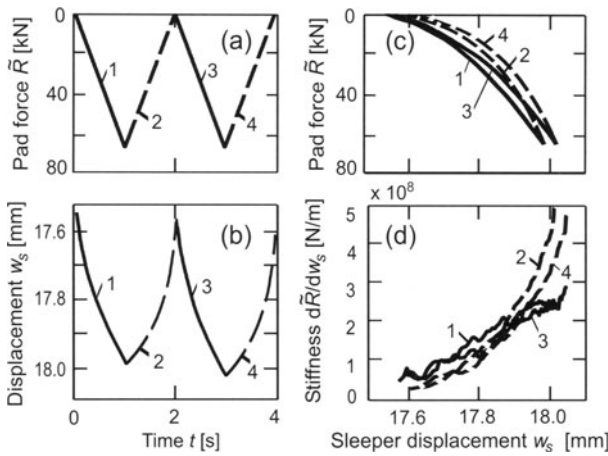
Fig. 16. Sleeper displacement during quasi-static loading

a function of sleeper displacement is presented in Fig. 14d. During loading the stiffness increases linearly, see curves 1 and 3. This is probably mainly caused by the quadratic contact law for the normal forces according to Eq. (2). In addition, also the formation of new contacts leads to a progressive stiffness. When the load has reached its maximum and starts to decrease, in the beginning the material has the tendency to remain in its stressed state, which means a jump in the stiffness curves. The reason for this phenomenon is the sticking of the stones — a back-sliding is only possible after the tangential contact forces have changed their direction and reached the amplitude $F_T = \mu F_N$. Very similar curves as those depicted in Fig. 14 have been measured in triaxial tests [3].

Kinematic Excitation In chapter 3 it is mentioned that an unfavourable superposition of travelling waves in the subsoil can cause track vibrations with high amplitude. Furthermore, strong vibrations of the ballast layer can be observed on bridges. How such an excitation influences the ballast settlement is demonstrated in Fig. 16. During the simulation runs the frequency is kept constant, while the amplitude is increased stepwise from $\hat{w}_f = 0.48\,\mathrm{mm}$ to $\hat{w}_f = 0.64\,\mathrm{mm}$ up to $\hat{w}_f = 0.8\,\mathrm{mm}$. These values correspond to maximum accelerations of $0.75g$, $1.0g$ and $1.25g$. The two curves, which have been determined for different constant pad forces $\tilde{R}$, show a drastic growth in the settlement rate when the acceleration of the wall particles reach the acceleration due to gravity g. This fluidisation effect has also been observed in experiments. The parameter studies may indicate that this effect can be reduced by prestressing the granular material ($\tilde{R} = 40\,\mathrm{kN}$). However, each of the two curves is just the mean of three program runs. Thus, much more simulations should be performed.
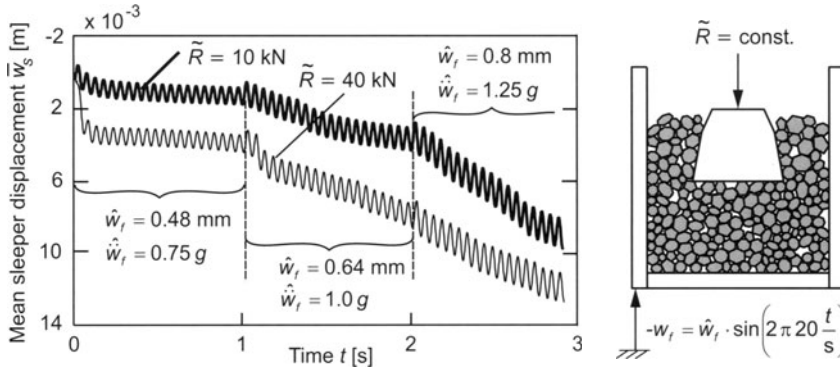
Fig. 17. Settlement behaviour of sleeper for kinematically excited ballast layer

6.3   Long-Term Behaviour

Decay of Stone Migration To prove the long-term stability of the ballast layer, simulations have been carried out with a high number of load cycles. The calculated sleeper displacement for six such simulations with about 700 load cycles is depicted in Fig. 17. As the friction coefficient is the most important parameter, three different values are used.
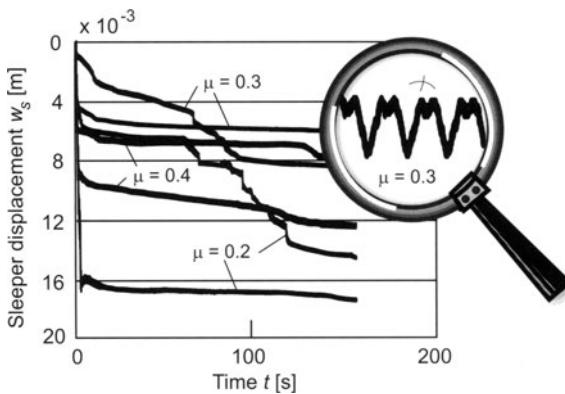
Fig. 18. Sleeper settlement for different friction values μ

In most cases the settlement decays rapidly. Nevertheless, after hundreds of load cycles still time segments with strong rearrangements inside the ballast layer occur. Although the smallest friction coefficient $\mu = 0.2$ leads to the highest settlement, the influence is lower than one would expect from Fig. 12. This may be a consequence of the fact that just two simulations for each long-term behaviour are not representative enough. Here again further research is needed.

Consideration of Abrasion A realistic calculation of the long-term behaviour requires a consideration of abrasion and stone breakage. Though, in principle the MD method allows for a description of these phenomena, the current knowledge of the underlying interrelationships is poor. Hence, in the following we introduce a very simple abrasion law, where the mass m and, therefore, also the cross section A of a stone decrease according to the formula

$$A\left(t+\Delta t\right)=\frac{m-k_w\cdot\tilde{F}}{m}\cdot A\left(t\right) \tag{9}$$

with $\tilde{F}$ as sum of all contact forces acting on a certain stone. The wear intensity is controlled by the wear factor k_w. Fig. 18 demonstrates the sleeper displacement for three different values of k_w and two different initial configurations.
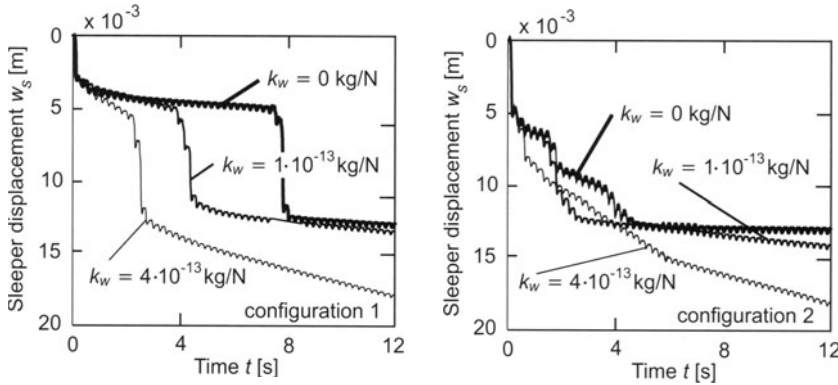
Fig. 19. Sleeper settlement for different wear coefficients k_w

For configuration 1, $k_w = 0\,\text{kg/N}$ gives a fast decay of the settlement. Only at $t = 8\,\text{s}$ a severe reorganisation takes place. If k_w is raised, as one would expect, this event is shifted to an earlier time point. Besides, the negative slope of the curves during phases without any noticeably stone migration increases as a consequence of the shrinking process.

With start configuration 2 it becomes obvious that a prognostication of the system behaviour is not simple. In Fig. 18b certain stone reorganisations occur even later due to the stone size reduction leading to intersections between the three curves.

7   Concluding Remarks

The results presented show that the MD method is applicable to the investigation of railway ballast. The stone layer forms a stable configuration in

respect to both static and dynamic loading, and this even in the case of free ballast surfaces or missing side walls, respectively. In addition, the system behaviour is strongly influenced by the initial conditions. Since e.g. the deviations in the sleeper settlement reach up to 60 % of the settlement's mean value, reliable statistic findings require a high number of program runs. In consequence, an enormous numerical effort is needed.

The boundary conditions and the choice of the contact law parameters play a less important role. Regarding the latter ones, even a wide variation of c_N, c_T, D_N or D_T does not lead to a substantial mutation of the system dynamics. The most relevant contact parameter is the friction coefficient.

Summarising, the calculation of contact forces, of kinetic and potential energies or of the ballast layer stiffness all yield plausible time dependencies, which are partly well-known from experiments. It is worth to be mentioned that some of the characteristics observed in the simulation — like on the one hand the jump in the quasi-static stiffness when the maximum compression is reached and the load drops or on the other hand the fluidisation during kinematic excitation — can hardly be reproduced on the base of an alternative continuum description. Regarding fields of application, the following interesting tasks can be investigated:

1. Reduction of ballast settlement. Influencing factors are e.g. axle load, train velocity, size of contact area between sleeper and ballast, thickness of ballast layer, stiffness of subsoil, ballast stone shape and coefficient of friction.
2. Optimisation of track components like sub ballast mats or elastic sleeper boots.
3. Optimisation of tampering processes.
4. Investigation of the effects of voids between sleeper and ballast.
5. Calculation of track shifting resistances.

An analysis of the long-term behaviour of railway ballast is an extremely difficult task. One possibility is to study the processes over just a few load cycles and to predict wear and stone breakage by evaluating the contact force network or the friction work. Otherwise, an extremely high number of train passings has to be simulated. Unfortunately, at the moment the contact model and the wear models are not sophisticated enough to provide reliable results.

Finally, some statements about an extension or improvement of the model shall be given: Certainly, the contact laws can be validated and refined by performing experiments with single stone-stone contacts or by using FEM tools. Alternatively, also a parameter identification via a comparison to triaxial or ballast box tests is conceivable. Nevertheless, the 2D-model can only provide qualitative results. For a 3D-model the computational power of current computers is not sufficient yet.

References

1. Bode, C., Hirschauer, R., Savidis, S. A. (2000) Three-dimensional time domain analysis of moving loads on railway tracks on layered soils. In: Chouw & Schmid (Editors): Wave 2000, Balkema, Rotterdam, 3–12
2. Cundall, P. A., Strack, O. D. L. (1979) A discrete numerical model for granular assemblies. Geotechnique **29**, 47–65
3. Guldenfels, R. (1996) Die Alterung von Bahnschotter aus bodenmechanischer Sicht. Veröffentlichung des Instituts für Geotechnik der ETH Zürich **206**, vdf Hochschulverlag Zürich
4. Knothe, K., Wu, Y. (1998) Receptance behaviour of railway track and subgrade. Archive of Applied Mechanics **68**, 457–470
5. Kruse, H., Popp, K., Krzyzynski, T. (1998) On steady state dynamics of railway tracks modelled as continuous periodic structures. Machine Dynamics Problems **20**, 149–166
6. Kruse, H., Popp, K. (2000) The influence of wave propagation in the subsoil on the train-track dynamics. In: Chouw & Schmid (Editors): Wave 2000, Balkema, Rotterdam, 171–184
7. Kruse, H., Popp, K. (2001) A modular algorithm for linear, periodic train-track models. Archive of Applied Mechanics **71**, 473–486
8. Kruse, H. (2002) Modellgestützte Untersuchung der Gleisdynamik und des Verhaltens von Eisenbahnschotter. Universität Hannover, Dissertation
9. Leykauf, G., Mattner, L. (1998) Moderne Fahrweg-Systeme. Eisenbahntechnische Rundschau **47**, 133–138
10. Matuttis, H.-G. (1998) Simulations of the pressure distribution under a two dimensional heap of polygonal particles. Granular Matter **1**(2), 83–91
11. Estrade i Panades, J.-M. (1989) Contribucion al Conocimiento del Macanismo de Deterioro de la Geometria de la Via por el Analisis del Comportamiento en Rotura de los Materiales que Forman la Capa de Balasto. Dissertation, Barcelona
12. Popp, K., Bogacz, R. (1984) Dynamik und Bewegungsstabilität von Zug-Gleis-Systemen. VDI-Berichte **510**. VDI Düsseldorf, 197–204
13. Schinner, A., Matuttis, H.-G. (1998) Internet: http://octopus.th.physik.uni-frankfurt.de/~schinner/granular/movies.shtml
14. Schünemann, A., Augustin, S., Huber, G., Gudehus, G. (2000) Sackungen des Schottergleises durch Zugüberfahrten. Schriftenreihe des Instituts für Grundbau und Bodenmechanik **32**, Uni-Bochum, 223–239
15. Selig, E. T. (1998) Ballast deformation: its causes and cures. Railway Track & Structures, May 1998, 25–31

The Dynamics of Railway Track and Subgrade with Respect to Deteriorated Sleeper Support

Michael Plenge and Rolf Lammering

Institute of Mechanics, University of the Federal Armed Forces Hamburg, Holstenhofweg 85, 22043 Hamburg, Germany

Abstract. Full scale laboratory experiments are performed in order to investigate the dynamic behavior of a ballasted track segment and its constituents. Since voids are often observed between sleepers and the underlying ballast, special emphasis is laid on investigations of partially unsupported sleepers. Detailed information on the dynamic displacement field is obtained by the holographic interferometry which is used beside conventional accelerometers to record the oscillations. It is shown that deviations from optimal coupling to the subgrade result in significant changes of the dynamic behavior. The results of the laboratory experiments are in good agreement with in-situ measurements. Elementary configurations are also corresponding well with numerical results.

1   Introduction

The current development of modern railway systems is particulary characterized by increasing travelling speeds, driving forces, and axle loads. From a technical point of view, this trend is requiring higher standards not only for each individual component but for the system as a whole. As a consequence, a careful overall design has to be guaranteed and especially the interaction between the vehicle, track, and subsoil has to be taken into account.

Typical damages of the track or its components are unequal ballast settlements, ballast damages, and as a result enlarged track positioning faults. The damage development process is initiated by small inhomogeneities or irregularities such as voided sleepers, non-uniformly compacted ballast, rail joints, or stiffness alterations, e.g. in bridge transition sections. Inhomogeneities and irregularities are causing track and subgrade vibrations during train passages giving rise to long term damages. Finally, they are responsible for increasing maintenance effort, cf. [6] and [7].

In the last 20 years, considerable theoretical and experimental work has been undertaken to understand the system behavior and to improve track components in order to solve the above mentioned problems. Dependent on the problem under consideration, various models have been developed for the entire system or its constituents covering a wide range of complexity, cf. [6] and [11]. It should be noted that beside the ballast especially voided sleepers cause a significant non-linear system behavior that has to be taken into account in a detailed analysis. Simple models consist of springs and dampers

connected in parallel, more sophisticated models use finite element or boundary element approximations in order to regard properly the influence of the ballast and the subgrade.

For non-railway applications, accurate models for the interaction of rigid and elastic foundations and the subsoil have been developed, e.g. by Gaul and Plenge [5], [4], and Wolf [16]. The interaction between adjacent rigid foundations through the subsoil has also been considered in [5]. A related problem was studied by Auersch [2] who determined subsoil properties from wave propagation data refering directly to the rail traffic.

Until now, only little work has been done that aims at the interaction between the subgrade and the railway track or its constituents, e.g. sleepers, adjacent sleepers, and systems of sleepers connected by rails. Auersch et al. [1] present numerical results of rigid sleepers on an elastic half-space under the assumption of a linear-elastic material behavior. The coupling between adjacent sleepers through the subsoil is considered experimentally by Rücker [14], and by Prange et al. [12]. They refer to in-situ experiments on tracks with removed rails in which the sleepers were excited. Recently, new numerical models have been developed in this field which are presented in accompanying papers within this book.

In order to evaluate theoretical and numerical models, laboratory experiments offer the advantage of higher accuracy compared with in-situ tests. This is the motivation for the series of full scale laboratory experiments presented in the following. The experimental set-ups cover various configurations from the single sleeper to the ballasted track segment with and without optimal coupling. Different kinds of excitations are applied to the measuring objects. Special attention is paid to voided sleepers which are considered as a long term damage as mentioned before. However, within this project the main concern is not the development of the voids but their impact on the short-term dynamics. The obtained data are compared to in-situ experiments as well as to numerical results.

For a pointwise capture of dynamic data, piezoelectric accelerometers and a laser vibrometer are used. Further insight is provided by means of the holographic interferometry which is used to present the entire displacement field. The frequency range is up to 1600 Hz, thus taking into account e.g. the relatively short wavelength irregularities of the rail and polygonalized wheels besides the regular vertical and lateral excitations.

2   Experimental Set-Up

The experimental investigations presented in this paper are performed in a laboratory in which an excavation of the dimensions 4.0m x 4.0m x 2.0m is filled with a test soil. Undisturbed soil is directly bordering the test soil without any walls so that almost any impedance difference is avoided. Therefore, even low frequency waves with a high wavelength may propagate without

reflections. The test soil is a mixture of sand and fine gravel with the properties shown in Table 1. This testbed allows for investigations on full-scale track segments consisting of up to five concrete sleepers B70W60 linked with rails UIC60 that are commonly used for ballasted German high-speed railway tracks, cf. Fig. 1.

Table 1. Properties of the ballasted test track

Object	Property	Data
rail UIC 60	mass per length	$\rho A = 60.34$ kg/m
	Young's modulus	E = 210,000 MPa
	Poisson's ratio	$\nu = 0.3$
	gemetrical moment of inertia y-y	$I_{yy} = 3,055$ cm^4
	geometrical moment of inertia x-x	$I_{xx} = 513$ cm^4
railpad Zw 900	static stiffness	$c_{stat} = 42.9$ kN/m
	dynamic stiffness (10 Hz)	$c_{dyn} = 91.4$ kN/m
sleeper	length	$l = 2600$ mm
	width	225 mm ¡ w ¡ 300 mm
	height	176 mm ¡ h ¡ 230 mm
	mass	m = 304 kg
	space between sleepers	$d_s = 600$ mm
ballast	density of ballast solids	$\rho_{bs} = 2,870$ kg/m^3
	density	$\rho_b = 1,750$ kg/m^3
	dynamic deformation modulus	E = 31.7 MN/m^2
soil	shear wave velocity	$c_s = 149$ m/s
	density	$\rho_s = 1,909$ kg/m^3
	Poisson's ratio	$\nu = 0.29$

For the loading of the track segment and its constituents, it is necessary to apply static as well as dynamic forces. Two pneumatic cylinders supported by a crossbeam generate a static load of 10 kN each. This load is about one tenth of the operating load, but our own experience (see Section 4.2) as well as investigations made e.g. by Wu and Thompson [17] show that a small static load influences the dynamic behavior of the system significantly but that a further loading causes only little modifications. The dynamic loads are applied by electrodynamic shakers, the force amplitude of which is amplified by a lever mechanism, as well as by rotating unbalanced weights.

The dynamic behavior of the structure under investigation is realized by piezoelectric accelerometers and by a laser vibrometer as well as by holographic interferometry. The accelerometers and the vibrometer allow for the collection of time signals at specific locations, whereas the holographic interferometry is an appropriate tool to measure the displacement field of the oscillating surface and the propagating wave at a definite time, cf. Fig. 1.

The basic principles of the holographic interferometry is outlined as follows. A beam of coherent light emitted by a laser is divided into two waves by a

beamsplitter. Each narrow ray bundle is then expanded by a lens or a system of lenses. The reference wave is directly guided onto a photosensitive plate of high resolution. The object wave is reflected by the surface of an opaque object and both waves interfere on the hologram plate in such a way that the variation of the relative phase from point to point is transformed into an interference pattern retaining the complete information of the object. The interference pattern is decoded by illuminating the hologram plate by the reference light and the virtual 3D-image of the object is obtained.

Fig. 1. Testbed with track segment and visualized laser beams

In order to record the displacement field, the holographic film is exposed twice at a very short time interval. During this period of time, the wave under consideration is propagating and the difference of the surface shapes is recorded on the film. Reconstruction of the hologram by reference light shows both images of the object at the same time. The object is covered by light and dark interference fringes that represent contour lines of the deformation. More details are given by Gaul and Plenge in [4].

3   Scope of Experimental Investigations

The test program starts with experimental investigations on a single concrete sleeper B70W60 which is widely used on German railway tracks in order to

identify the eigenfrequencies and eigenmodes. Various boundary conditions are realized: free - free, perfectly coupled to the subsoil, and voided sleepers, i.e. sleepers with hovering ends. A hovering end of a sleeper arises when a gap between the soil and the sleeper grows during operation. The gap may be built either at one end of the sleeper or at both ends. In the latter case, the contact between the sleeper and the subsoil exists only in the middle of the sleeper. Both situations are depicted in Fig. 2. The frequency range under consideration is up to 800 Hz.

Fig. 2. Different kinds of voided sleepers

Beside the sleepers, the dynamic bahavior of the rail under vertical and lateral excitation is investigated. Subsequently, two adjacent sleepers are under consideration which are coupled through the subsoil. Again, different boundary conditions and voided sleepers are studied. The investigations on non-ballasted configurations end up with an entire track segment consisting of five sleepers and the connecting rails. The results which have been presented in an earlier paper by Lammering and Plenge [8] were very supporting for the understanding and interpretation of ballasted track results.

The ballasted track segment is the final configuration within the series of experiments. Voided sleepers are again in the center of the investigations. The load is applied vertically and laterally. These experiments serve as the basis for comparisons with in-situ experiments and numerical investigations. The dynamics of a railway track is significantly influenced by railpads that are mounted between the rails and the sleepers. Therefore, they have been investigated as individual structural members and additionally, their impact on the entire track has been studied.

4   Identification of the Concrete Sleeper B70W60

4.1   Sleeper with Free - Free Boundary Conditions

In order to realize the free - free boundary conditions and to exclude the influence of the support in the following investigations, the concrete sleeper is rotated with respect to its axis by 90^0 and suspended by two ropes. The excitation is performed horizontally (i.e. in the rotated z - direction) at one end by an electrodynamic shaker. The positions of the shaker and selected accelerometers as well as the frequency response function are presented in Fig. 3.

From there it can be seen that the inertance function is characterized by narrow peaks of bending modes with respect to the x - axis. In Fig. 3, the
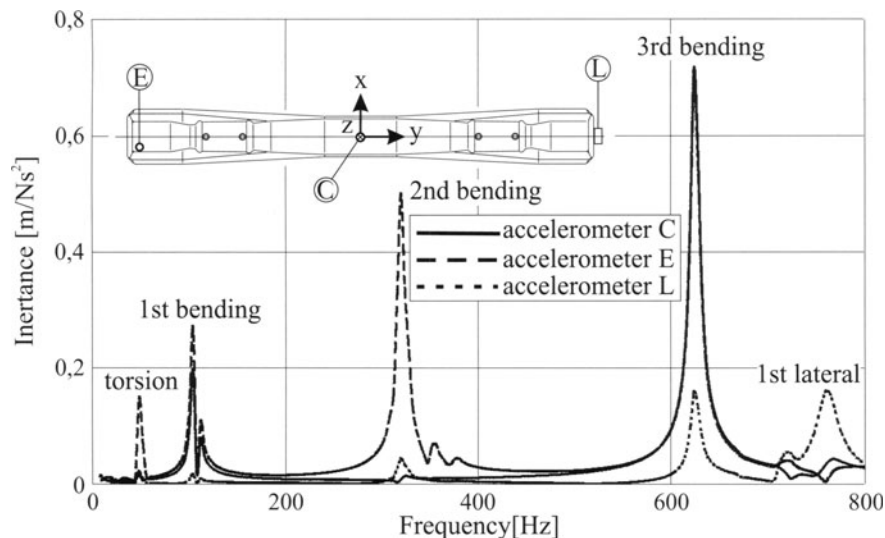
Fig. 3. Frequency response function of the rotated and horizontally excited sleeper

peaks are labelled with the corresponding modes. Beyond that, particular accelerometers give rise to additional peaks, e.g. at 52.5 Hz and at 760 Hz. At 52.5 Hz, the torsional mode is excited as a result of an imperfection in the experimental set-up and observed by the accelerometer E that is mounted opposite to the shaker position with an offset from the center axis, cf. Fig. 3. At about 760 Hz, the first longitudinal mode of the sleeper occurs which would result in a lateral vibration of the track.

The experiment was repeated with the same sleeper which again was suspended but not rotated. Again, the sleeper is excited horizontally, i.e. in the x - direction. The frequency response function is similar to that shown in Fig. 3. The first three bending modes, however, are shifted about 9 to 15% towards higher frequencies, cf. [8].

4.2   Ideal Contact Between Sleeper and Subsoil

Next, the sleeper is laid on the test soil and perfectly coupled with the ground. Optimal coupling conditions are achieved by a plaster on which the sleeper is placed after it is spread on the soil in the contact area. Various tests were performed in which the static load and the direction of excitation (vertical, lateral and tilt) were varied among other. From the frequency response functions those frequencies were identified by which the system is substantially affected. At these frequencies, the displacement field of the soil and the sleeper is recorded by use of the holographic interferometry. The influence of the static load that is applied at the railseat area is shown in Fig. 4. The load is increased in steps of 2.5 kN up to 10 kN. It is obvious that the first load

step results in a stronger coupling of the sleeper to the soil, so that especially
the third bending mode at about 632 Hz is less developed. On the other hand,
the frequency response function is almost not affected between 250 Hz and
550 Hz. Below 250 Hz, any increase of the static load produces stiffening of
the system. These results are in good agreement with investigations of Wu
and Thompson [17] on a ballasted track. They report that an increase of the
static load from 1.2 kN to 11 kN gives rise to the stiffness (factor 2.4) at a
dynamic excitation at 50 Hz, but that further doubling of the load increases
the stiffness by 5% only.
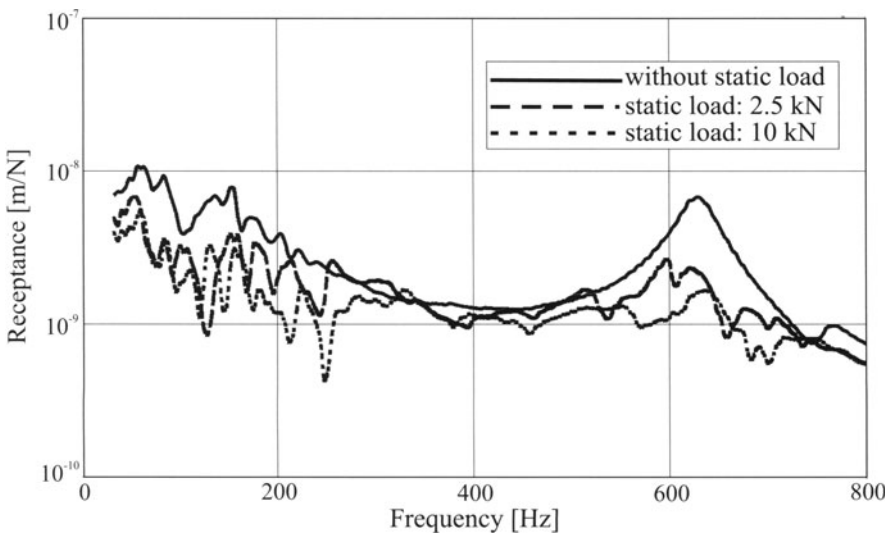
Fig. 4. Influence of the static load on the frequency response function of an ideally
coupled sleeper

From the experiments above, the frequency of 148 Hz is selected for further
investigations focussing on the oscillations of the sleeper and the wave prop-
agation in the soil. Fig. 5 shows the displacement field of the sleeper and
the soil for two different phases that are shifted by 180^0. The vibration of
the sleeper is dominated by the first bending mode. Furthermore, it is visible
that the wave propagation in the soil starts at both contact areas between the
sleeper and soil. At 148 Hz, the wavelength is about 1 m, so there is a shear
wave velocity of about 149 m/s in the soil. This result is in good agreement
with seismic measurements. It should be noted that the interferogram was
subdivided into two parts before the evaluation is started since the edge of
the sleeper or shadows can influence the evaluation algorithm. Therefore, the
displacement field is evaluated separately for the sleeper and the soil.
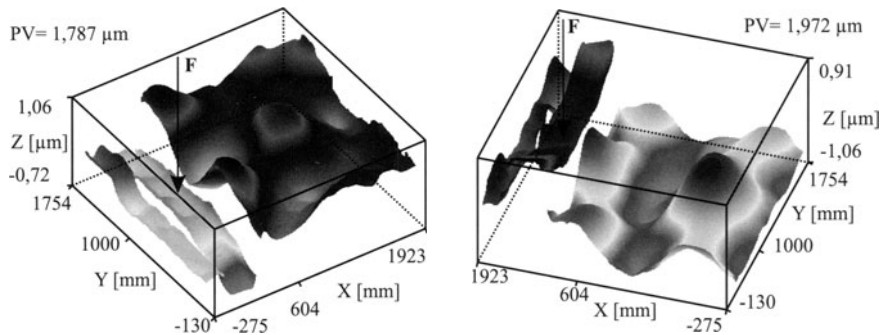
Fig. 5. Dynamic displacement fields of the soil and the sleeper in the case of ideal coupling and excitation in the sleeper center at 148 Hz. 180^0 phase difference between both figures

4.3   Partially Unsupported Sleepers due to Voids

The ideal contact between the sleeper and the subsoil is often lost during operation, and a gap between the sleeper and the soil is built up either at one end or at both ends. This long-term damage is rebuilt in the laboratory. The receptances of the ideally coupled and voided sleepers are compared
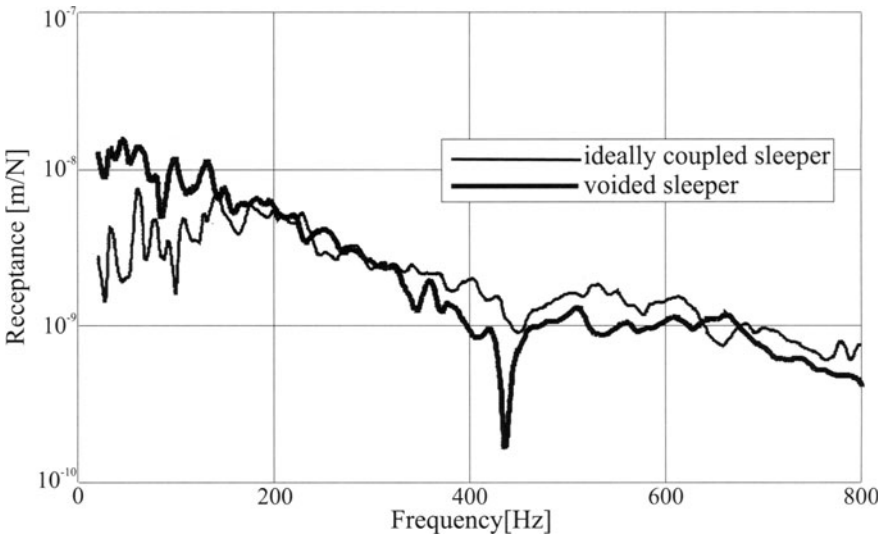
Fig. 6. Comparison of the receptances of the ideally coupled sleeper and a voided sleeper. 10 kN static load at each rail

in Fig. 6. For this purpose, a static load of 10 kN is applied at the rail positions and a dynamic load is introduced at the center. The signal of the

accelerometer attached at the upper side of the sleeper end is shown as well as the corresponding signal from the ideally coupled sleeper. It is visible that both curves are in good agreement beyond 150 Hz. With the exception of a small frequency band at app. 440 Hz at which a kind of anti-resonance appears, the difference does not exceed the factor 2. On the other hand, the voided sleeper behaves much softer in the frequency range below 150 Hz.

Fig. 7. Dynamic displacement fields of the soil and the sleeper with central coupling and excitation at one railseat at 52.5 Hz. 90^0 phase difference between the four figures

Because of a small resonance amplification at 52.5 Hz in the frequency response function the holographic measurement technique is applied to evaluate the dedicated mode of vibration. The vertical excitation force acts on one railseat without any static load. In Fig. 7 the dynamic displacement fields of sleeper and surrounding soil surface are shown for four oscillations states, of which the phase positions are shifted by 90^0 from one figure to the next. From the left figures with a time delay of half a vibration period the coupling conditions can clearly be detected, because the surface waves propagate from the contact area in the middle of the sleeper. The sleeper ends are obviously without contact to the soil. Because of the low frequency the displacement

field at soil surface looks similar to the static case. The sleeper itself shows only small deformations and tilts around the contact area. The pictures on the right hand side depict the bending oscillation of the sleeper which is superposed by a torsional mode of the free sleeper ends. At this particular trigger time the vertical displacements at the soil surface are very small.

5   Structural Dynamics of a Track Segment

The following experiments are performed on a track segment consisting of five sleepers and two connecting rails. The track segment is bedded either on sand or in ballast under various coupling conditions. In order to measure the frequency response behavior of the system it is dynamically loaded by use of electrodynamic shakers up to 1.6 kHz. The phase angle between the loads on each rail is varied as well as the level of the static load. Nine accelerometers and two force transducers are mounted in various patterns on the rails and the sleepers in order to record the oscillations. One pattern was used to prove the symmetry of the response when the system is loaded symmetrically. The rail vibration is observed by another pattern, in which the accelerometers are mainly mounted on the top of the rail. Here, two places of excitation are chosen: on the sleeper and between the sleepers.

In the following, some results are presented which were obtained from the ballasted track segment. The main task of the project is to evaluate the influence of the coupling conditions between sleeper and ballast. Therefore after completion of the experiments with optimal coupling an artificial void was implemented at the right hand side of the sleeper in the middle of the track segment, cf. Fig. 1. The opposite part of the same sleeper as well as the other four sleepers are perfectly coupled. For both configurations the whole system is dynamically loaded by two forces acting vertically on both rails between the second and third sleeper. The frequency response functions in Fig. 8 are obtained when both forces are acting to the rails 180^0 out of phase. The upper diagram shows the transfer functions for the perfectly coupled sleeper end with respect to the coupling conditions. The void at the opposite side lead to a higher receptance in a wide frequency range. However, the lower figure shows clearly the minor stiffness of the sleeper end above the void in the frequency range up to 120 Hz. Beyond this frequency the bedding stiffness changes scarcely . On the other hand, it can also be seen in comparison of both diagrams that even in the case of perfect coupling the mean value of the receptances differ at both sleeper ends in the lower frequency range. This aspect will be discussed in detail in Section 6.

The dynamic behavior of a track is strongly influenced by elastic railpads. Before installation two types of pads Zw 900 were tested in a material testing device. In Fig. 9 the quasi-static load-displacement curves show that their elastic behavior is very different from each other in spite of the fact that the secant modulus that serves a parameter in the condition of supply of the
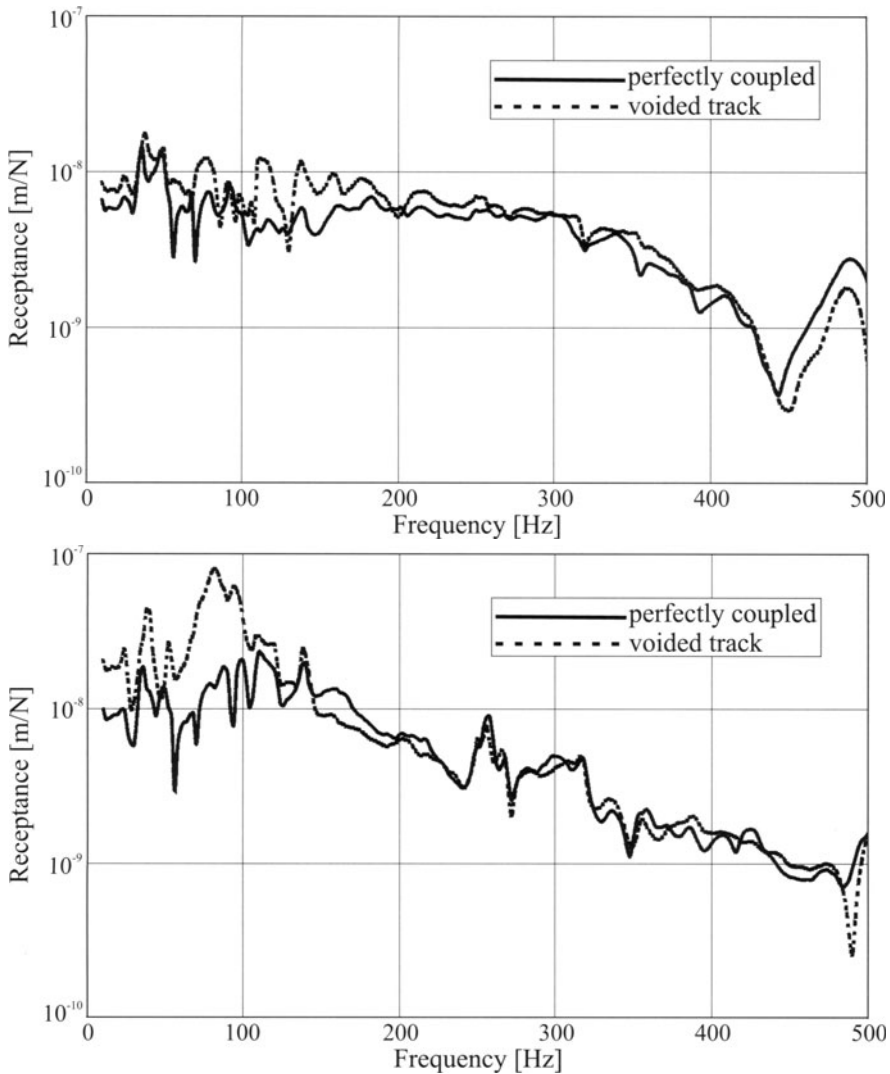
Fig. 8. Receptances of a track segment with a voided sleeper. Upper figure: left sleeper end, lower figure: right sleeper end

railway companies are the same. As the loads on the track in the experiments are comparatively small it is advantageous to install the railpad type 2 in the test track because of the nearly linear behavior in compression.

The influence of the railpads Zw 900 is investigated for a perfectly coupled, ballasted track segment under lateral excitation in counter-phase at the rail-heads above the sleeper in the middle of the track segment. The structural response is also measured at the railheads opposite to the loading points.
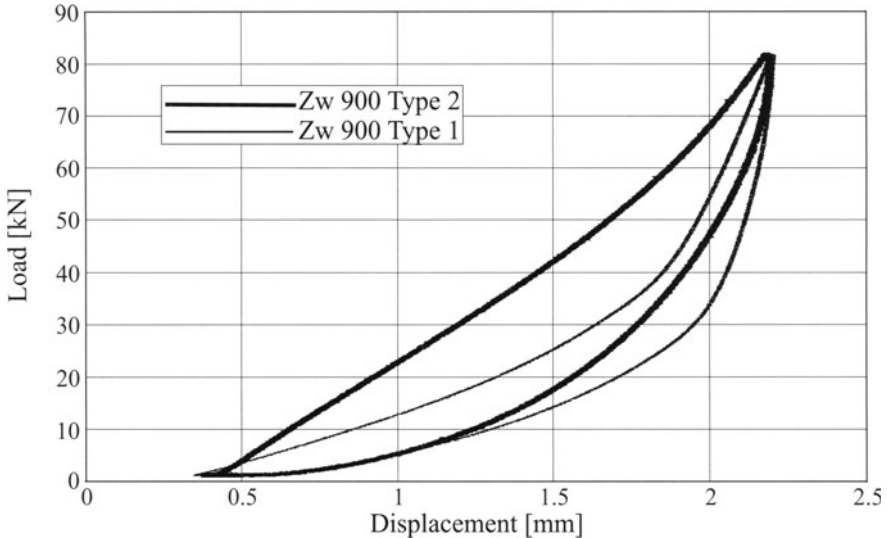
Fig. 9. Railpad load-displacement curves

The frequency response function without (above) and with (below) railpads for the left and right rail are shown in Fig. 10. In the case without pads the influence of ballast and subgrade can be seen in the frequency range up to 250 Hz because of the different bedding stiffness under the sleeper ends which was already detected for vertical loading (Fig. 8). After installation of the pads this effect does not occur any more. Refering to the excitation the odd bending modes of the rails show up very distinctly in the receptance function.

By means of the holographic interferometry dynamic displacement fields of a perfectly coupled track segment on sand (Fig. 11 above) and ballast (Fig. 11 below) were measured under vertical dynamic loads at 103 Hz in phase and in counter-phase. The results are shown for three sleepers and the spaces between them. If the track is directly bedded on the soil surface and excited in phase the oscillations of the sleeper are associated with the first bending mode. The adjacent sleepers twist because of the coupling through the rails and the soil. Furthermore, it is clearly visible that the waves spread out from the sleeper in the middle of the track segment. On the right-hand side of Fig. 11, where the situation is shown when the forces are 180⁰ out of phase, the displacement field of the sleeper and the surrounding soil is quite different. Loads in counter-phase induce a rocking oscillation of the whole track. Also the waves propagating from the coupling areas move in counter-phase to the oscillation of the sleeper because the excitation frequency is higher than the rocking mode of the track. The diagram below shows the displacement fields for the ballasted track under the same loads generated by rotating unbalanced weights. Because of the rough surface of the ballast
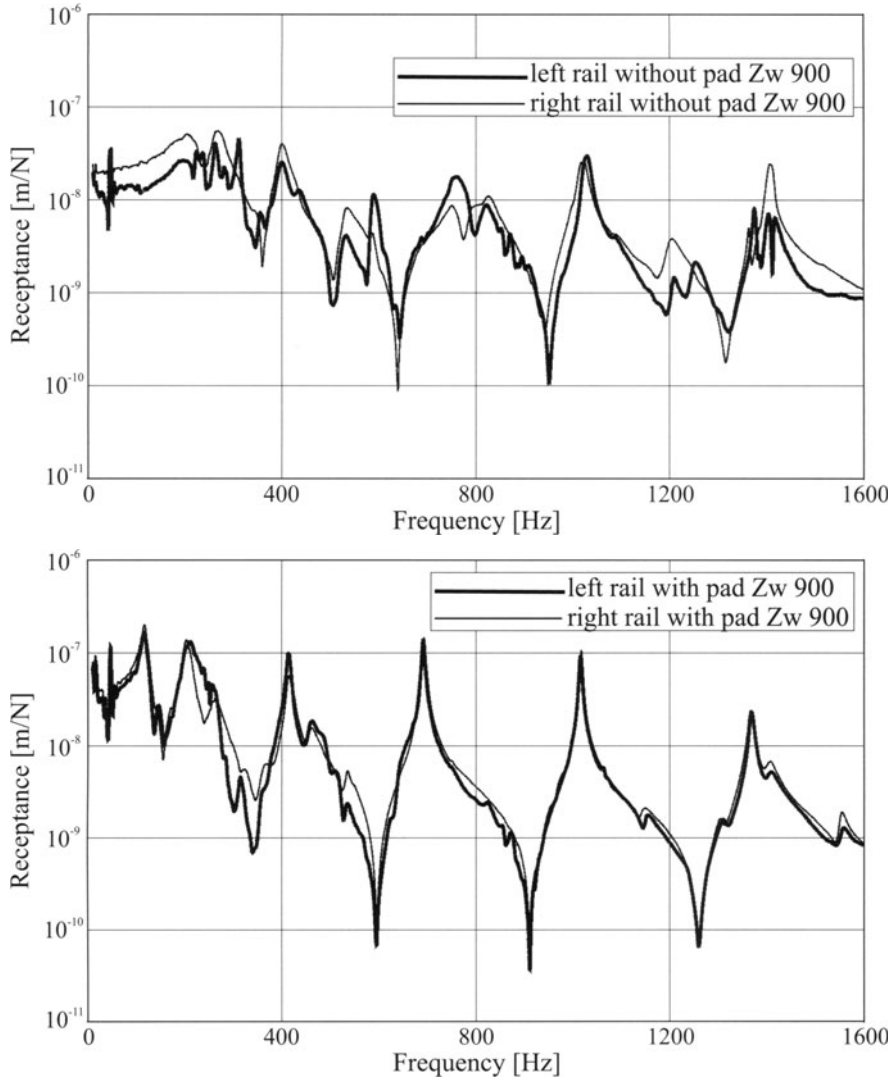
Fig. 10. Railpad influence on the track receptance at lateral load. See text for further explanation

and small movements of its particles relatively to each other an evaluation of the displacement fields was not possible for the ballasted area. The bending and rocking oscillations of the sleeper and additionally of the rails are very similar to those of the track on soil. This fact proves that the investigations on the dynamic behavior of a track on sand as a intermediate step are very important for a better understanding of the behavior of a ballasted track.
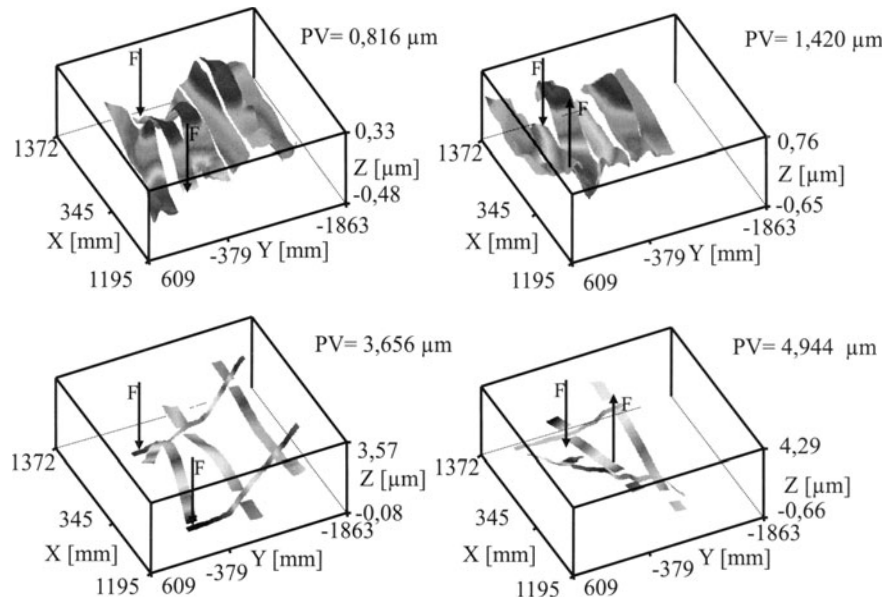
Fig. 11. Dynamic displacement fields of the track segment and the surrounding soil. See text for further explanation

Another result of the holographic interferometry is presented in Fig. 12 for a laterally loaded ballasted track segment at an excitation frequency of 95 Hz. The dynamic displacement fields of the sleepers are shown for two oscillation states delayed by 180^0 . The rocking mode caused by the lateral loads in counter-phase is overlayed by a strong bending motion of the sleeper under excitation. This bending oscillation is caused by the artificially deteriorated sleeper suport under one side of the sleeper described above. The evaluation of such effects is a useful tool for the identification of voids.
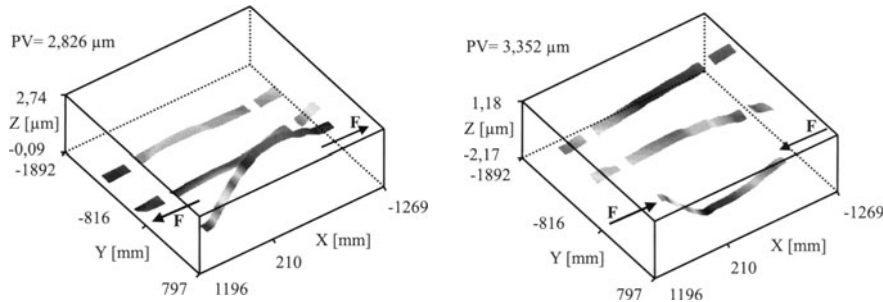
Fig. 12. Dynamic displacement fields of the track segment at lateral loading

6   Comparison Between Laboratory and In-Situ Experiments

In order to ensure that the insights obtained by the laboratory experiments are of any value for practical problems, it is necessary to compare the results from laboratory experiments with data obtained by measurements at ICE-high-speed tracks under live loads, cf. [9]. For these in-situ experiments on five successively arranged sleepers accelerometers were mounted at the rail-head and sleeper-end in vertical direction. The first measurements were carried out at a newly built track loaded by an ICE-train passing with the velocity of 252 km/h. The signals were analysed as third octave spectra in such a way
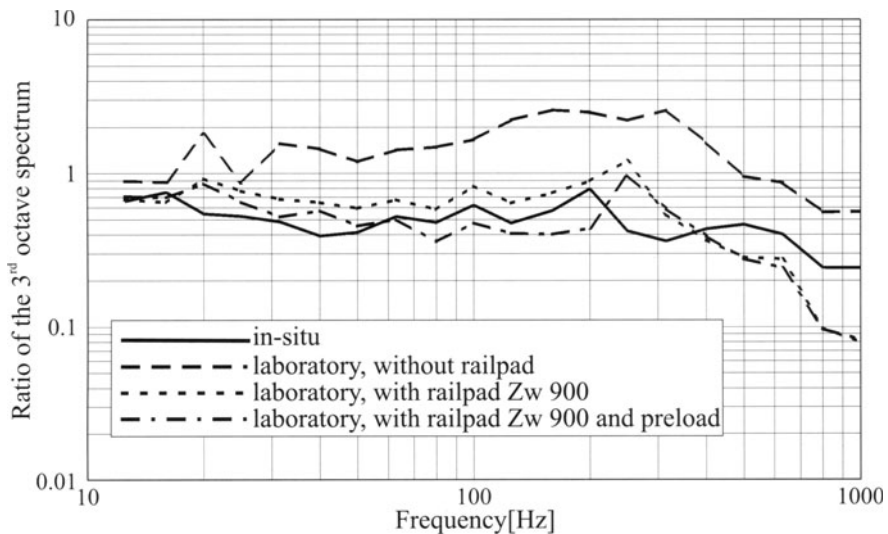
Fig. 13. Ratio of the 3rd octave velocity spectra between sleeper and railhead

that only wagons are enclosed in the time slot; so-called wagon-spectra. In order to eliminate the excitation transfer functions were calculated dividing the velocity spectrum at sleeper-end and the velocity spectrum at rail-head. The first comparison of the in-situ results with laboratory measurements, shown in Fig. 13, was insufficient because the test track was not equipped with railpads, cf. [10]. After installation of the same type of pads like those in the ICE-tracks the agreement was much better in the whole frequency range. In a third step a static load of 10 kN was applied to each rail. As a result the curves match still better even in the lower frequency range.

The in-situ measurements were repeated one and a half year later after the track had been loaded by approximately 7 Mio. loadtonnes. In this case the ICE-train velocity amounts to 250 km/h. By means of these measurements a gap between sleeper 2 and the ballast was detected [9], because, as shown
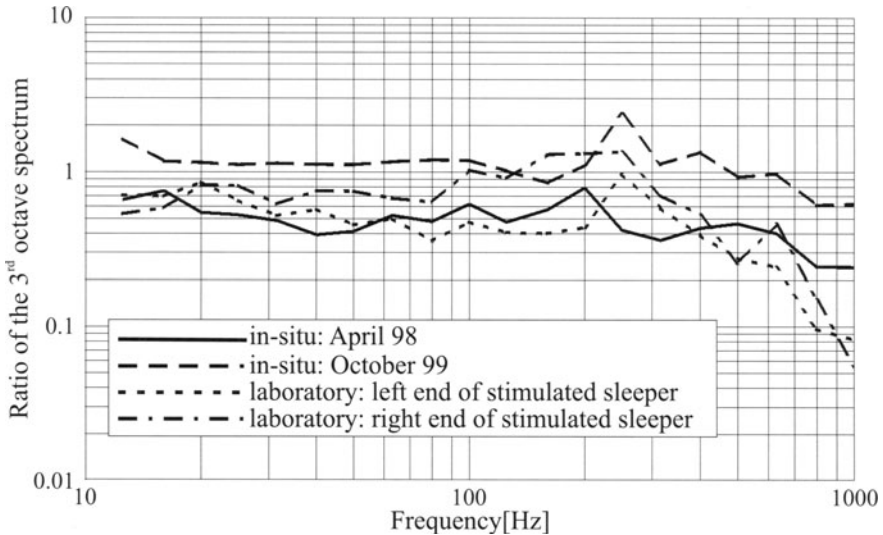
Fig. 14. Ratio of the 3rd octave velocity spectra between sleeper and railhead for a track with a voided sleeper

in Fig. 14, the sleeper support is much softer compared to the originally results in Fig. 13. Similar results were obtained in the laboratory for the opposite ends of the sleeper under excitation although a perfect coupling was expected. Within the whole frequency range from 20 Hz to 800 Hz the stiffness on one sleeper end is significantly diminished. In spite of very good working conditions in the laboratory the building company did not manage to produce a perfectly ballasted track bed in this area. On the other hand it is shown that sleepers with softer bedding conditions are undoubtedly identifiable with the used measurement technique.

7   Comparison Between Experimental and Numerical Investigations

An important task of the described project is to collect reliable experimental data obtained in a well known experimental environment for theoretical model verification. Therefore a close collaboration takes place with those work groups of the DFG-priority-program which develop theoretical models to calculate the dynamic behaviour of the track-subgrade-system. The measured data of the geometrical and material properties of the ballasted track shown in Table 1 are the basis for the numerical calculations. The comparison of experimental and theoretical results start with the consideration of elementary systems. First a single sleeper on sand is investigated which is vertically excited at its center. If the sleeper is modelled as a rigid beam, the calculated frequency response function matches with the curve from ex-
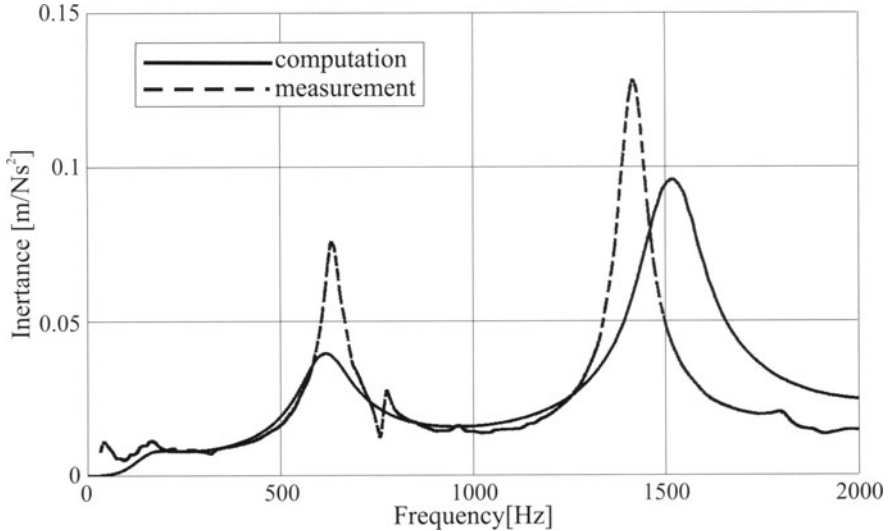
Fig. 15. Comparison between numerical and experimental results in the frequency domain

periment only in an average, because the bending and torsional oscillation modes of the sleeper cannot be realized by this kind of model. In contrast the inertance function at the end of an elastic sleeper on sand calculated by means of the thin layer method [15] coincides much better with the measured result even in the frequency range up to 2000 Hz, cf. Fig. 15. The resonance frequencies of the odd bending eigenmodes of the sleeper match quite well. The measured amplitudes in the range of the resonances are generally larger. Probably this increase is caused by a reduction of the geometrical damping due to imperfections of the coupling conditions in the contact area. Deviations in the low frequency range are ascribed to the torsional oscillation mode of the sleeper at 52.5 Hz. Additionally, the lateral eigenmode is observed only in the measured frequency response function at 780 Hz.

First results for a single elastic sleeper on sand in the time domain are presented in [3] by Firuziaan and v. Estorff. The agreement of the measured and calculated displacement-time graphs at sleeper end respectively to the first bending oscillation mode is quite well up to 20 ms; but the decayed oscillation in the experiment is not covered by the numerical simulation, cf. Fig. 16. Furthermore, the overlayed oscillation of the third bending eigenmode does not occur in the numerical results. For other configurations, especially for the ballasted track, a lot of calculations exist, but the agreement with the experiments is currently not sufficient. Further work is necessary to identify the reasons in order to achieve a better compliance.
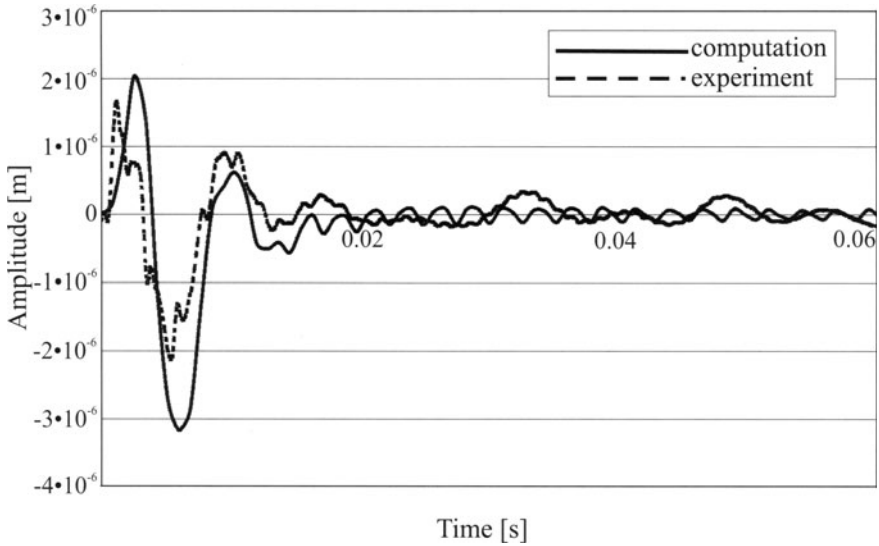
Fig. 16. Comparison between numerical and experimental results in the time domain

8   Conclusions and Outlook

In this project, the dynamic behavior of a single sleeper, two adjacent sleepers and a track segment under vertical and lateral loads is under consideration. Special emphasis is laid on the influence of voided sleepers on the short-term dynamics.

Investigations on a single sleeper show that a voided sleeper gives significant rise to the receptance at frequencies up to 150 Hz. Even in the case of vertical excitation, these boundary conditions allow for rigid body modes and torsional modes, which are not observed when the sleeper is perfectly coupled to the ground. At frequencies beyond 150 Hz, the dynamical behavior does not change significantly.

Investigations of two adjacent sleepers and a track segment have shown that some of the characteristics which have been observed on a single sleeper, play also an important role in the behavior of more complex structures. E.g., the same bending modes can be found not only in the frequency response functions of a single sleeper, but also in a system consisting of two adjacent sleepers and of a track segment.

A static load results in a stronger coupling of the track to the ground and in stiffening of the system. Our own experiments confirm the observations made by Wu and Thompson [17] who note that the main effect is already visible at comparatively small loads.

Since the ideal coupling to the ground is inevitably lost during track operation, track models could be improved by taking into account the results of

these laboratory experiments and by incorporating the effect of voided sleepers. Likewise, the increase of the receptance below 120 Hz are worth to be considered in the analysis of vehicle - track interaction. Deteriorated sleeper supports are clearly detectable by structural dynamic parameters regardless of the kind of excitation.

The holographic interferometry is an appropriate measuring technique to get comprehensive understanding of the interaction between railway track and subgrade.

In spite of precise specifications by the railway companies, the load displacement curves of railpads are not unique. On the other hand, their influence on the dynamic behavior of the track system is significant.

The results from the laboratory experiments are in good agreement with in-situ measurements from high-speed tracks under live loads. This agreement is obtained when the laboratory and the in-situ tracks are built in the same way, i.e. with respect to the rails, the sleepers, the railpads, and the ballast, and when a static load of 10 kN is applied to each rail in the laboratory experiments. Thus, the transferability of the laboratory results to practical problems is guaranteed.

Numerical results obtained from different theoretical models map the dynamic behavior of a single sleeper on sand satisfactorily in the frequency as well as in the time domain. However, numerical and experimental results of the entire track segment are in poor agreement as of yet. Here, further research is required.

Within this project, important knowledge on the dynamical behavior of well positioned and voided tracks has been collected. Further investigations are required to get a broader database for structural health monitoring.

9   Acknowledgement

References

1. Auersch, L., Rohrmann, R., Rücker, W., Said, S. (1987) Fahrzeug-Fahrweg-Untergrund-Umgebung: theoretische, rechen- und meßtechnische Untersuchungen, VDI Berichte Nr. 635, Dynamik fortschrittlicher Bahnsysteme, 59-87
2. Auersch, L. (1994) Wave Propagation in Layered Soils: Theoretical Solution in Wavenumber Domain and Experimental Results of Hammer and Traffic Excitation. Journal of Sound and Vibration, 173 (2), 233-264
3. Firuziaan, M., v. Estorff, O. (2002) Simulation of the dynamic behavior of bedding-foundation-soil in the time domain. In: Popp K., Schiehlen W. (Eds.)

System Dynamics and Long-Term Behaviour of Railway Vehicles, Track and Subgrade, Lecture Notes in Applied Mechanics, Springer-Verlag

4. Gaul, L., Plenge, M. (1992) Progress in 3-d BE-calculations and optoelectronic measurement of soil-structure-interaction. Elsevier Applied Science, London, Advanced Dynamic Analysis by BEM, Developments in Boundary Element Methods - 7, Eds. Banerjee, P.K., Kobayashi, S., Chapter 10, 353-403

5. Gaul, L., Plenge, M. (1993) Baugrundeinflüsse auf das Schwingungsverhalten von Fundamenten, VDI Berichte Nr. 1082, Dämpfung und Nichtlinearität, 195-220

6. Knothe, K. (2001) Gleisdynamik. Verlag fr Architektur und technische Wissenschaften GmbH, Berlin

7. Knothe, K., Grassie, S. L. (1993) Modelling of Railway Track and Vehicle/Track Interaction at High Frequencies", Vehicle Systems Dynamics, 22, 209-262

8. Lammering, R., Plenge, M. (2000) Investigations on railway tracks with special emphasis on partially unsupported sleepers due to voids. Engng. Trans., 48, 293-307

9. Müller-Boruttau, F. H., Ningler, M. (2000) Daten von Messungen. Aktennotiz N189012 v. 31.5.2000

10. Plenge, M., Lammering, R. (2000) Änderung des Schwingungsverhaltens des Gleisoberbaus unter Vertikal- und Lateralanregung infolge von Gleislagefehlern. VDI-Berichte Nr. 1568, Dynamik von Fahrzeug und Fahrweg, 41-64

11. Popp, K., Schiehlen, W. (1993) Fahrzeugdynamik. Eine Einführung in die Dynamik des Systems Fahrzeug - Fahrweg. B. G. Teubner, Stuttgart

12. Prange, B., Huber, G., Triantafyllidis, Th. (1982) Dynamisches Modell des einzelnen Gründungskörpers, 2. Technischer Bericht zum Forschungsvorhaben des BMFT TV 78 150 Teil B: Dynamisches Rückwirkungsmodell des Gleisoberbaus: Feldmessung und FE-Modelle, korrigierte Fassung

13. Ripke, B., Knothe, K. (1995) Simulation of High Frequency Vehicle-Track Interactions", Vehicle Systems Dynamics Supplement 24, 72-85

14. Rücker, W. (1980) Messungen der Reaktionen mehrerer Schwellen, die über den Untergrund gekoppelt sind, infolge harmonischer Erregung einer starren Schwelle in verschiedenen Erregermoden, Technischer Bericht zum Meilenstein D4 des Vorhabens Rad/Schiene, Förderungskennzeichen TV 7817 2, Bundesanstalt für Materialprüfung, Berlin

15. Savidis, S., Bode, C., Hirschauer, R., Schepers, W. (2002) Dynamic interaction between railway tracks and layered subsoils. In: Popp K., Schiehlen W. (Eds.) System Dynamics and Long-Term Behaviour of Railway Vehicles, Track and Subgrade, Lecture Notes in Applied Mechanics, Springer-Verlag

16. Wolf, J. P. (1985) Dynamic Soil - Structure Interaction. Prentice Hall, Englewood Cliffs

17. Wu, T. X., Thompson, D. J. (1999) The Effects of Local Preload on the Foundation Stiffness and Vertical Vibration of Railway Track, Journal of Sound and Vibration, 219(5), 881-904

Part V:

Subgrade Dynamics

Numerical Model and Laboratory Tests on Settlement of Ballast Track

Sven Augustin[1], Gerd Gudehus[1], Gerhard Huber[1], and
Andreas Schünemann[1]

Institute of Soil and Rock Mechanics, University of Karlsruhe,
Engler-Bunte-Ring 14, 76131 Karlsruhe, Germany

Abstract. The long-term behaviour of ballast track was investigated. Tests on a railway model circular circuit showed troughs and saddles fixed in place. The ballast under badly bedded sleepers undergoes greater vertical deformation than under well bedded sleepers. Cyclic loading tests of ballast showed a decisive influence of stress minimum on the permanent deformation. In cyclic element tests the state variables of the material vary within each cycle around their mean values after initial cycles. The mechanical response – e.g. stiffness – changes strongly. Simulations of these element tests with a hypoplastic material law modelled the main features of the mechanical behaviour of ballast correctly including increase of stiffness, decrease of hysteresis and influence of stress minimum. Based on the former results a numerical model of the track was developed showing also the growth of height differences. Initial state has strong influence on long term behaviour. Irregular track geometry with characteristic wave lengths develop. Variation of density leads to higher permanent and to higher differential changes of ballast height.

1 Questions

The long-term behaviour of ballast track is closely related to the questions:

- Why do track irregularities build up?
- How do they develop?
- How can they be reduced or avoided?
- How can the ballast material be described by constitutive models?

Understanding of the long term behavior is essential for getting answers to these questions. This is a basis for reducing life-cycle-costs and increasing quality and availability of the track.

Model and cyclic element tests have been carried out to observe the material and system behaviour in detail. A hypoplastic material law was used for numerical simulations of element tests and to model a railway track.

2 Irregular Settlements of Model Railway Tracks

2.1 Model Circular Track

To investigate the development of track-bed-waves model tests were performed using sugar grains as model ballast.
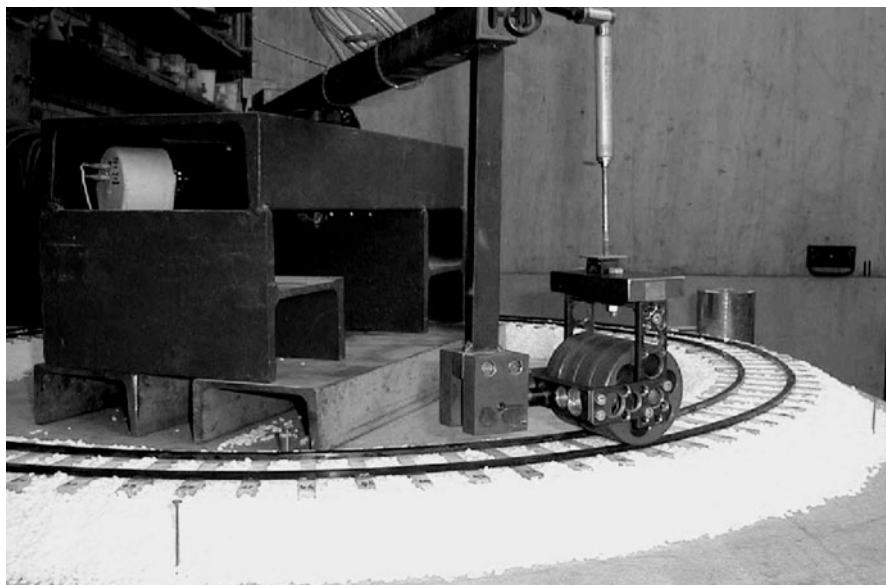
Fig. 1. Circular track on sugar ballast with roller

Figure 1 shows part of the track circuit with a diameter of 1 m on a rigid base. The sleepers are made of wood, the rails of PVC, the ballast is modelled with sugar grains with a grain size of 2-4 mm. The grains are similar to real ballast gravel with regard to their surface roughness and angularity. Fracturing of the sugar grains occurs with the equivalent model forces, approximately similar to grain breakage observed in situ. Track loading is simulated using a roller with a diameter of 80 mm. The roller is driven from the centre of the circuit and is constructed as a trailer to allow for free vertical motion.

Data Acquisition. Rail displacements due to the passing roller were measured with a linear variable differential transformer (LVDT). The reference position of the roller was measured using an inductive proximity switch. After each 100 circle passages the roller was supplied with the displacement gauge recording the vertical position during the passage around the circle. The difference between two data sets yielded the differential vertical displacement due to 100 passages. Furthermore, the elevation of grain surface beneath the sleepers was indirectly quantified. Therefore the vertical displacement of each sleeper during loading using a weight was measured. This procedure was performed before and after the experiment. Results of measurements are shown in Fig. 3.

Objectives. The model was developed in order to understand mechanisms of settlements in railway tracks due to train passages. Model tests were carried

out with allowance for ballast properties. Material parameters, in particular grain hardness and friction angle, were adapted to the model scale. Thus, abrasion and breakage of the ballast grains, as observed in situ, were achieved. Furthermore, the influence of increasing train speed was investigated: three tests were carried out with velocities of 0.5, 1.2 and 1.8 m/s. Another objective was the influence of initial imperfections within the track bed and their alterations due to subsequent train passages. In particular, the influence of initial density and voids beneath the sleepers was studied in detail.
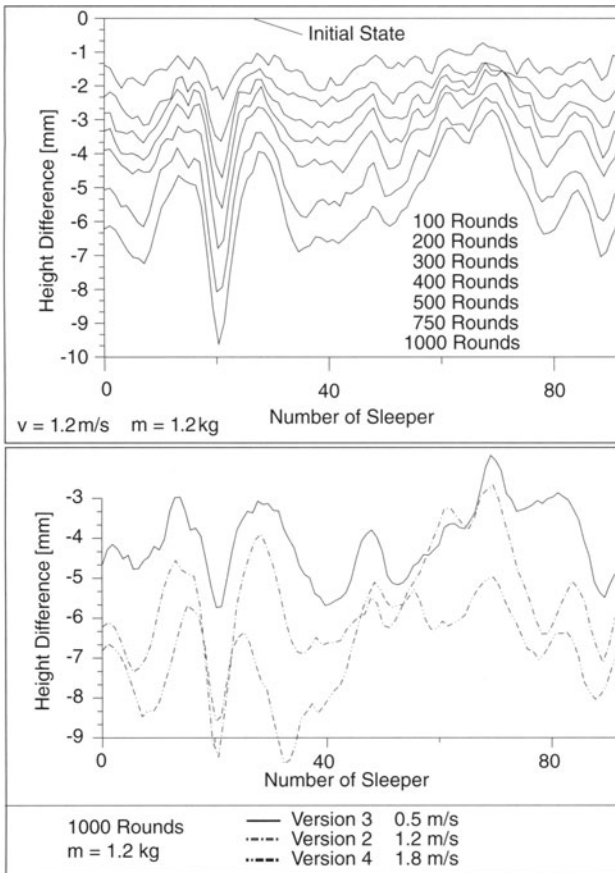
Fig. 2. Development of model track bed waves

Test Results. Figure 2 (above) shows the development of vertical displacements. The maximum displacement at sleeper number 22 was produced by an intentional defect in the rail.

Other waves originated from small irregularities, e.g. variation of ballast-density or badly bedded sleepers, which lose contact to the ballast after unloading.

It was found that once a wave developed, settlements quickly increase. This development of track-bed-waves is induced by the vertical acceleration of the roller, producing inertial forces. Figure 2 (below) shows the vertical displacements after 1000 passages for three different roller velocities. The displacements depend on the velocity with a maximum average settlement at the highest velocity. Disregarding sleeper 22 (defect in rail) the waves evolved at different locations. During the tests these positions were identified as those of badly bedded sleepers.
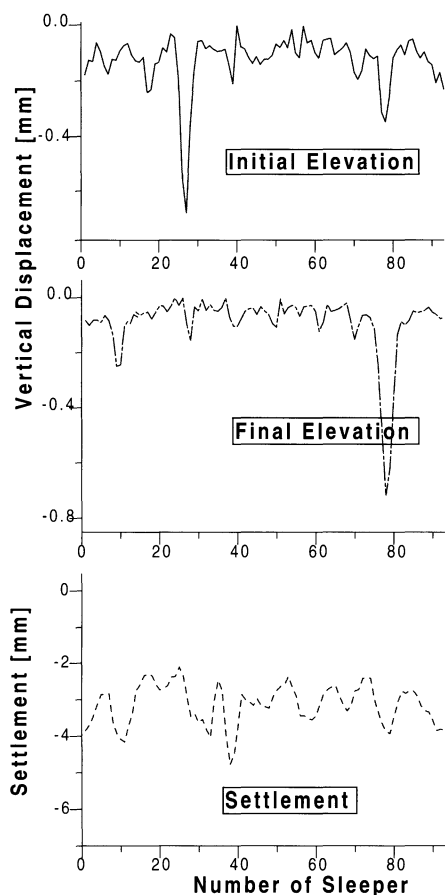
Fig. 3. Initial and final elevation, track settlement

Figure 3 shows the initial elevation, the final elevation and the settlement after 1500 rollings of the track bed. Two locations of badly lying sleepers can be recognized: sleeper 26 and 80. At sleeper 26 the void beneath the sleeper disappears, whereas the subsidence at sleeper 80 increases. Consequently there must be some difference between these two positions. The two following mechanisms are conceivable:
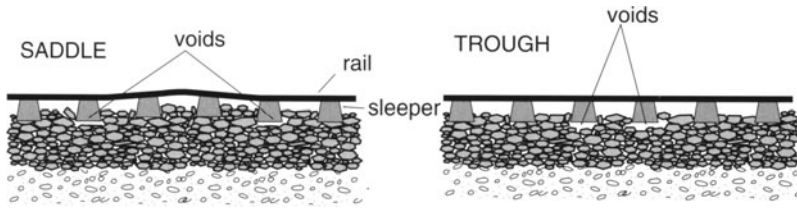
Fig. 4. Variation of void location beneath sleepers influences track bedding

If a sleeper is in a higher position than desired, the sleepers in front and behind lose contact to the ballast (Fig. 4). After a few passages the saddle will straighten out. This happened at sleeper 26. On the contrary a trough had formed around sleeper 80. At this location non-uniform settlements increase with each passage.

2.2   Ballast Loading Test

In another experiment the behaviour of the sleeper on ballast during cycling loading was investigated. Figure 5 shows the experimental setup: a cross-shaped footing made of concrete was used instead of sleeper; thus lateral tilting was minimized. The cross was loaded by a servo-controlled hydraulic actuator equipped with a load cell.

Objectives. The purpose was to investigate the influence of stress conditions on the ballast behaviour. Minimum stress beneath the sleepers varied between 0 kPa for badly placed sleepers and 10 kPa for well placed sleepers. During a train passage the stress increases up to 200 kPa depending on train load.

Test Results. The results are shown in Fig. 6. The ordinate at the right side shows the stress amplitude during the cycles.
The influence of minimum stress is clearly visible. With decreasing minimum stress an increasing settlement of the track bed could be observed.

2.3   Summary of Model Tests

Model tests simulating the accumulation of track displacements have been carried out.

Fig. 5. Cross on ballast
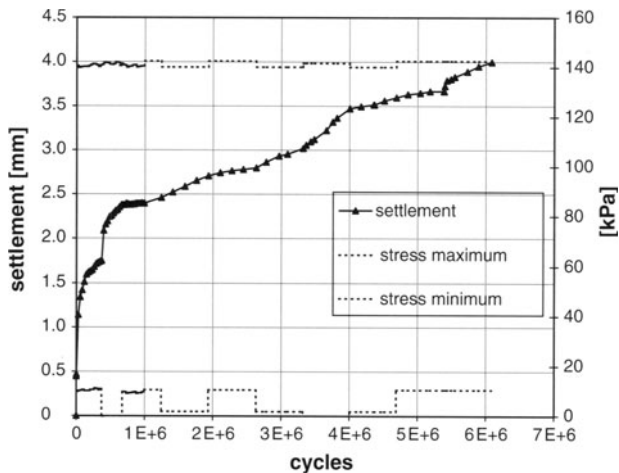
Fig. 6. Settlement of a cross on ballast during cycling loading (cf. Fig. 5)

The importance of initial conditions was clearified: Small irregularities such as inaccurately positioned sleepers or density imperfections in the ballast influenced the evolution of vertical displacements in the track.

This model helped to find a procedure that limits and reduces the development of track bed waves. It emphasizes a maintenance procedure regulating the rail position without influencing the ballast density or the geometry of the superstructure, and includes preventive measures to minimize the development of new voids beneath the sleepers.

3 Long Term Behaviour of Railroad Ballast in Laboratory Tests and Numerical Simulation

3.1 Cyclic Element Tests

Element tests are aimed to the study of stress-strain behaviour of granular materials, and to the calibration of material laws. All state variables, in particular stress and void ratio, are assumed to be uniform over the sample at each time (homogeneous state). In cyclic element tests the boundary conditions are cyclic.
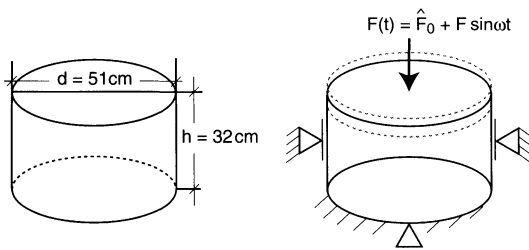
Fig. 7. Cyclic oedometer test: geometry of the sample and boundary conditions

Cyclic Oedometer Test. The material behaviour of railroad ballast has been investigated by large scale element tests. Figure 7 shows the boundary conditions of a cyclic oedometer test and the geometry of the ballast sample which has been prepared with medium initial density. Vertical compression increases linearly over the logarithm of number of cycles between the cycle no. 3000 and 768000, see Fig. 8.

The material experiences only small permanent deformation although it is far from maximum density. The permanent deformation $\varepsilon_{11,p}$, the reversible deformation $\varepsilon_{11,e}$ and the dissipation of energy by hysteresis W_h are listed in Table 1 for the 100th and the 768000th cycle. Both cycles are depicted in Fig. 9. As expected the tangential stiffness grows with the number of cycles.

Table 1. Comparison of the 100th and the 768000th cycle

cycle number	$\varepsilon_{11,p}$	$\varepsilon_{11,e}$	W_h
100	$1.1 \cdot 10^{-5}$	$6.5 \cdot 10^{-4}$	$9 \cdot 10^{-3}$ kNm/m^3
768000	$4.5 \cdot 10^{-10}$	$3.0 \cdot 10^{-4}$	$4 \cdot 10^{-3}$ kNm/m^3
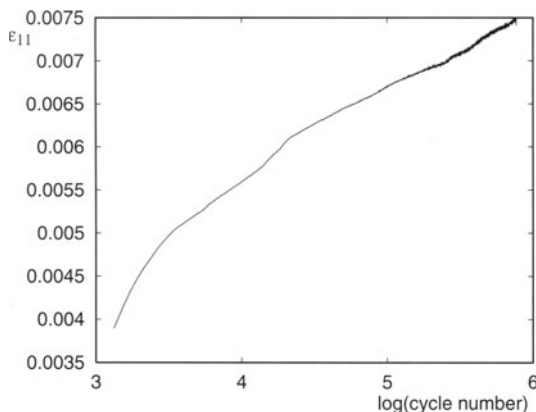
Fig. 8. Accumulated vertical deformations in a cyclic oedometer test
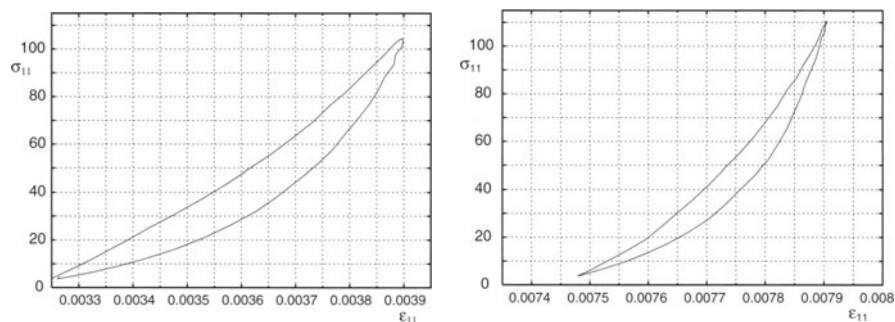
Fig. 9. Stress-strain curves of the 100th and the 768000th cycle

The dramatic reduction of $\varepsilon_{11,p}$ corresponds to the reduction of $\varepsilon_{11,e}$ and W_h by only a factor of 2.2 and 2.25 resp. So in spite of a nearly vanishing plastic deformation a considerable amount of energy dissipation is observed.

Cyclic Triaxial Test. The boundary conditions of a cyclic triaxial test and the geometry of the used ballast sample are shown in Fig. 10.

Due to allowed lateral strains the material is now able to compensate for vertical compression by horizontal extension. Even dilatancy can take place. Considering the accumulation of volume change a dense material dilates first and then very slowly compresses again. However all these effects take place in a very narrow range, see Fig. 11.

Despite the smaller number of cycles the triaxial sample exhibits more accumulated vertical deformation than the sample in the oedometer test. For further cycles the accumulated deformation is even underlinear.

Figure 12 shows the stress-strain curves of selected cycles. The observations for the oedometer test concerning $\varepsilon_{11,p}$, $\varepsilon_{11,e}$ and W_h hold also in the triaxial case, especially, the tangential stiffness increases considerably. The stiffness

$F(t) = F_0 + \hat{F} \sin\omega t$

$\sigma_3 = \text{const.}$

$h = 78 \text{ cm}$

$d = 80 \text{ cm}$

$\sigma_{11,min} = 25 \text{ kPa}$

$\sigma_{11,max} = 150 \text{ kPa}$

$\sigma_{33} = 25 \text{ kPa}$
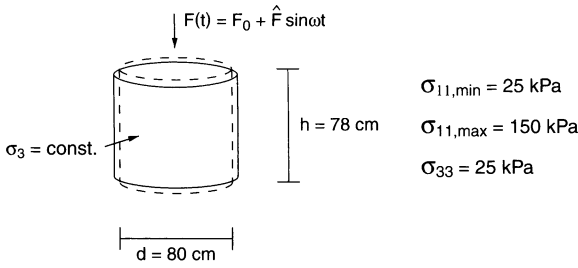
Fig. 10. Cyclic triaxial test: geometry of the sample and boundary conditions
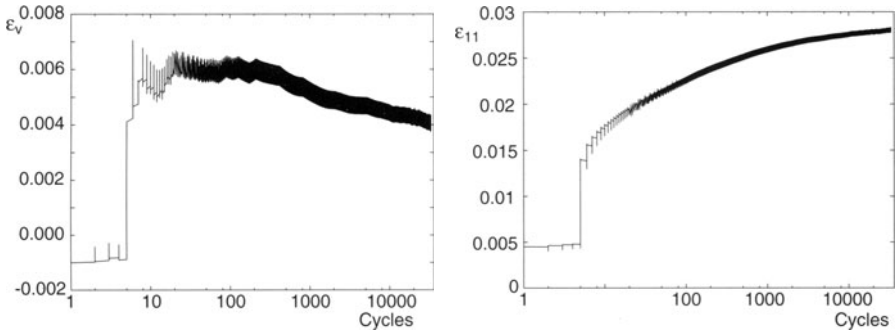
Fig. 11. Accumulated volumetric and vertical deformations in a cyclic triaxial test

varies strongly within each cycle. The values $\Delta\sigma_{11}/\Delta\varepsilon_{11}$ for cycle 30000 range from about 100MN/m^2 to 1100MN/m^2. Similar results showing this type of nonlinearity are found in laboratory and field tests ([6], [4], [3]).

The shape of the first cycles after starting the test looks quite different. The loading branch within a cycle undergoes a more significant change of the shape than within the unloading branch, Fig. 13, left. After cycle 47200 (Fig. 13, right) the ballast sample was compressed with $\dot{\varepsilon}_{11} = const.$ reaching a maximum stress $\sigma_{11} \approx 270 \text{kN/m}^2$ and then again cyclic loading was applied. The changes of the curve shapes during these cycles are also more pronounced in the loading branch than in the unloading branch of each cycle. After about 100 further cycles the shape of the loops becomes similar to cycles prior to monotonic compression.

Conclusions for the Cyclic Element Tests. In both cyclic tests the samples show a considerable change in the material behaviour with increasing number of cycles. In the triaxial case this is accompanied by an oscillation of the vertical stress σ_{11} and a nearly constant void ratio e. The reasons of this apparent contradiction may be abrasive processes and plastic deformations at the grain contacts. The same holds for the oedometer test, even though lateral stress is not known in this case. Further investigations should clarify these
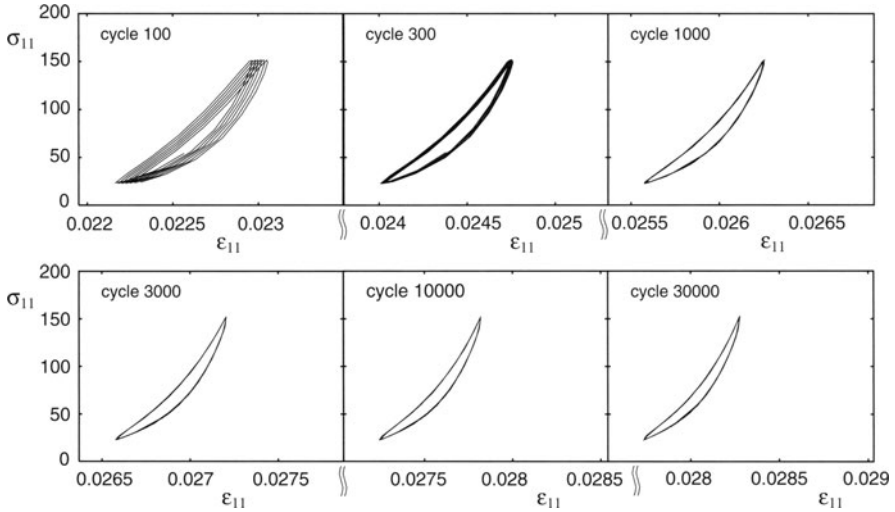
Fig. 12. Selected cycles of a cyclic triaxial test
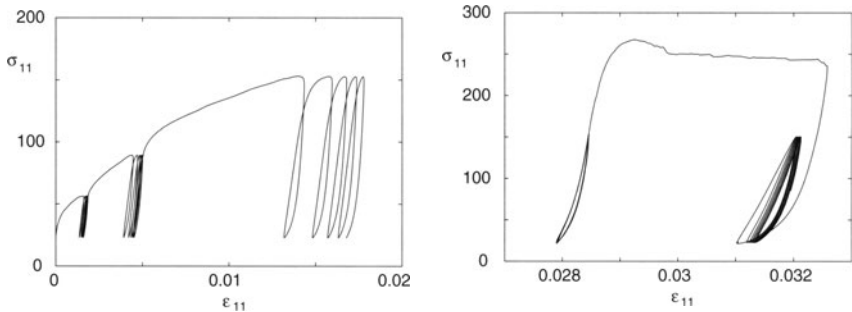
Fig. 13. Selected cycles of a cyclic triaxial test: first cycles (left), cycles (47180-48200) before and after peak (48201-48207)(right)

phenomena and related questions, e.g. whether additional state variables are involved.

3.2 Numerical Simulation of Cyclic Element Tests

Compared to the numerical simulation of monotonous deformation processes the calculation of material behaviour in the case of alternating deformation adds some further difficulties like strain history, abrasion processes and long term behavior. We present a constitutive law, some simulation results and discuss them with regard to the test results.

The Hypoplastic Constitutive Law. The stress-strain behaviour of a non-cohesive granular material can be expressed as a relationship between the objective stress rate $\overset{\circ}{\mathbf{T}}$ and the strain rate $\mathbf{D}$, written as

$$\overset{\circ}{\mathbf{T}} = \mathbf{h}\,(\mathbf{T}, e, \boldsymbol{\delta}; \mathbf{D}) \tag{1}$$

with three state variables: Cauchy stress tensor $\mathbf{T}$, void ratio e and the tensor of intergranular strain $\boldsymbol{\delta}$, which stores recent deformation history. A representation of $\mathbf{h}$ is given in [7]. [2,5] show that the hypoplastic equations have an objective mathematical and physical foundation. They are apt to model the material behaviour in a wide range of stress and density.

Cyclic Oedometer Test. The hypoplastic law has been used to simulate a cyclic oedometer test with 100000 cycles and the boundary and initial conditions of the laboratory test. Figure 14 shows the 100th and the 100000th cycle in a stress-strain representation.
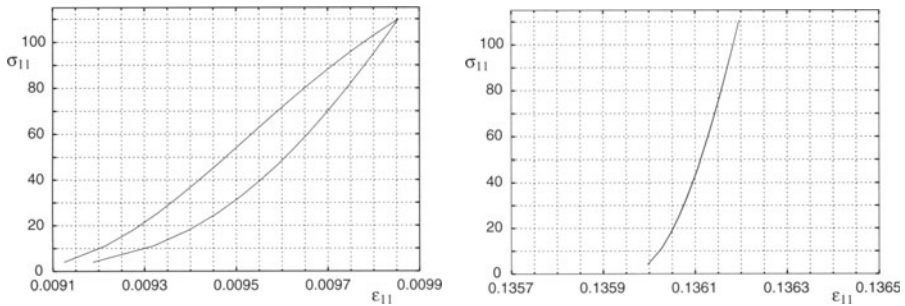
Fig. 14. Stress-strain diagrams of the 100th and the 100000th cycle (simulated by hypoplasticity)

We see that the qualitative properties of the outlined above material behaviour are well reproduced by the simulations: Extreme reduction of permanent deformation is accompanied by moderate reduction of reversible deformation and by an increase of tangential stiffnesses.

The curve of accumulated deformation is depicted in Fig. 15, which disagrees partially with the experimental results. The accumulation of the vertical deformation evolves slower in the first cycles and faster in the last, but it shows a linear characteristic in the logarithmic scale. This means that the pattern of logarithmic deformation is well reproduced by the hypoplastic law. As already mentioned the material behaviour in the experiments changes substantially but the state variables oscillate around mean values or remain almost constant. In order to obtain from calculations a considerable change of material
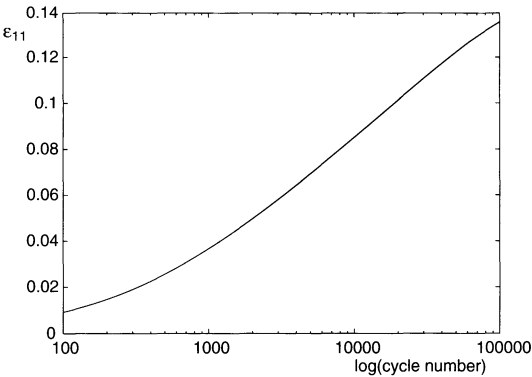
Fig. 15. Accumulated vertical deformations (simulated by hypoplasticity)

response, one needs a considerable change of the state variables. In the case presented here this happens by contractancy i.e. the increase of density, so that in comparison to the experiments we get too large permanent deformations. New concepts have to be developed, e.g. by an introduction of a new state variable, in order to get the quantitative side of the material behaviour in the cyclic case.
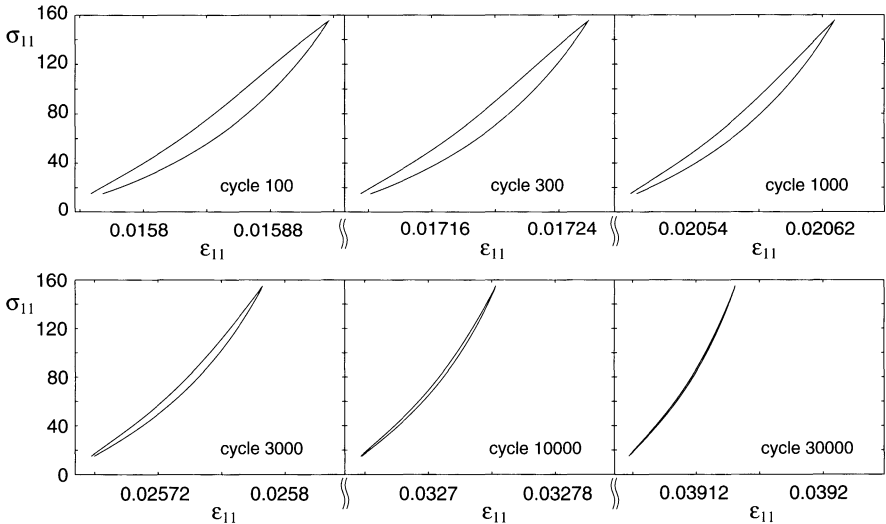
Fig. 16. Selected cycles of a numerically simulated cyclic triaxial test

Cyclic Triaxial Test. For a numerical simulation of a cyclic triaxial test we consider the selected cycles according to Fig. 12. In comparison to the

laboratory test we see in Fig. 16 the same material behaviour concerning the evolution of tangential stiffness, $\varepsilon_{11,p}$ and $\varepsilon_{11,e}$. The magnitude of accumulated deformation does not match for the reasons mentioned above, although we find that the material behaviour is well modelled qualitatively.

Conclusions from the Simulations of Cyclic Element Tests. The simulation results show that the hypoplastic law is able to predict the main features of material behaviour of ballast in the case of alternating stress/strain paths. The quantitative side has to be captured by an improved hypoplastic model.

4   A Numerical Track-Train Model

4.1   The Model of the Track

A numerical model that describes repeated passages of a vehicle over a rail-track-subsoil system has been developed. Figure 17 depicts the components of this system:
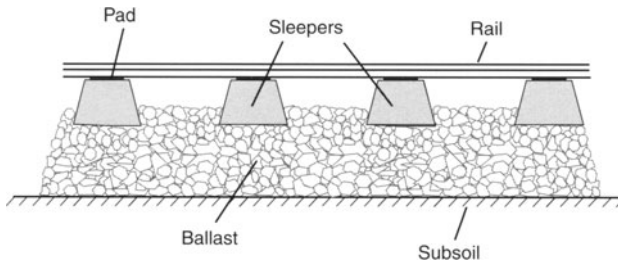
Fig. 17. Numerical model of a railway track

- the rail as Timoshenko beam with periodic boundary conditions,
- the pad as elastic spring,
- the sleepers as rigid bodies,
- the ballast as hypoplastic material with intergranular strain,
- the subsoil as elastic material with 3D-wave propagation and energy radiation.

The ballast body is divided in biaxial elements, one element per sleeper. Figure 18 shows the boundary conditions of such a ballast element. The state variables are assumed to be homogeneous in each ballast element. The considered length of the track is $36\,\mathrm{m}$ which corresponds to 60 sleepers. The driving system is represented by a standard two mass spring damper system, see Fig. 19.
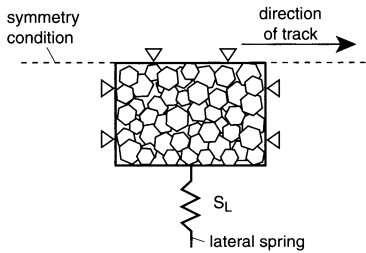
Fig. 18. Top view and boundary conditions of a biaxial ballast element

$d_s = 2.2 \cdot 10^4$ Ns/m
$c_s = 6.6 \cdot 10^5$ N/m
$m_2 = 9.5$ t
$m_1 = 0.5$ t
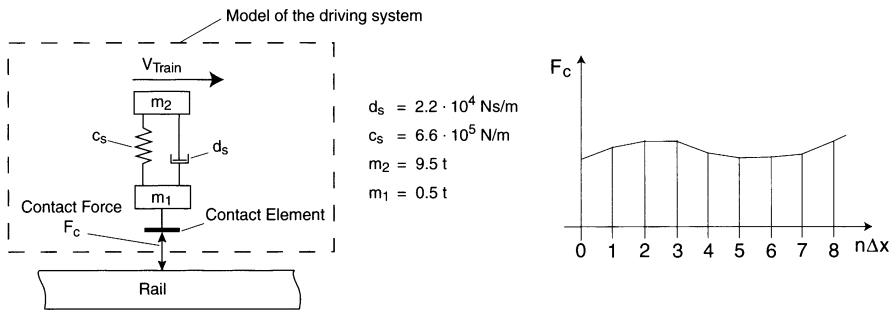
Fig. 19. The driving system and exemplaric evolution of contact forces

The contact force F_c between driving system and rail is determined by iterations until force equilibrium of the whole track-vehicle system is achieved. This is necessary due to acceleration forces acting on the driving system. Loss of contact is possible, but does not happen for velocities below 350 km/h. The problem of contact loss is more important in the case of the interface between sleepers and ballast. In the reality it is well known that a considerable number – mostly over 50% – of the sleepers are badly bedded or do not touch the ballast at all. Figure 20 shows such a situation, which is covered by the proposed model.
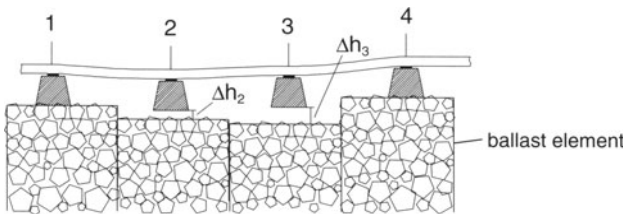
Fig. 20. Contact loss of sleepers: Δh is a measure of the troughs

4.2   Troughs as Reasons of Additional Permanent Deformation

In a trough the ballast in the unloaded state (no vehicle present) has no contact to the sleeper, so the minimum stress in the ballast is very small. Under a well bedded sleeper the minimum stress in the ballast is higher due to the load of the sleeper and the rail weight. This is more pronounced, if the neighbouring sleepers have no contact to the ballast, they are 'hanging' beneath the rail and increase the minimum stress in the saddle. In section 2.2 model tests have shown that accumulated vertical deformation by repeated loading is significantly increased when the stress minimum is reduced. This explains the positive feedback of the troughs in the prototype.
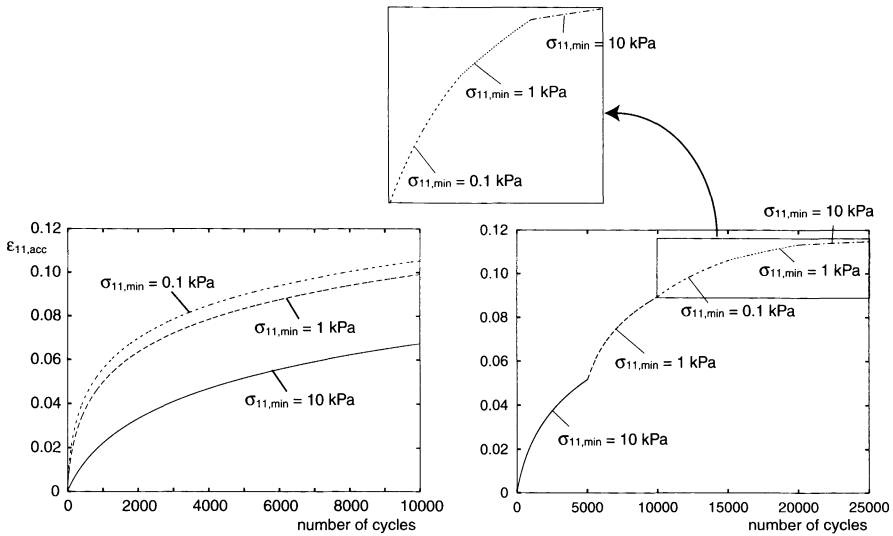
Fig. 21. $\varepsilon_{11,acc}$ for different $\sigma_{11,min}$

Figure 21 shows that in numerical simulation between cyclic oedometer tests the hypoplastic law gives the same dependence of accumulated deformation $\varepsilon_{11,acc}$ and stress minimum $\sigma_{11,min}$: The smaller $\sigma_{11,min}$ the higher $\varepsilon_{11,acc}$. Therefore we should expect a positive feedback in the simulation results of the track model as observed in the prototype.

Figure 21 provides also another important information. As can already be seen in the measurements of the ballast loading test, see Fig. 6, more than half of the accumulated deformation takes place during the first cycles. In Fig 21, right, we see two effects: First, the largest accumulated deformation results from the first cycles. Second, reducing $\sigma_{11,min}$ produces an increase of rate of accumulated deformation; increasing $\sigma_{11,min}$ reduces this process. This observation suggests to overcome as soon as possible initial cycles in order to come through large deformations and not to return to such a state,

which results in large settlements. During tamping for maintenance purposes the ballast comes to such a sensitive state. Consequently this method should be replaced.

4.3   Simulation Results

Originally the numerical simulation aimed to model long term behaviour. It came out that time incrementation is needed to be so fine that this demand seems to be unreachable at present. Thus the model has been used to carry out a parameter study with 10000 passages of the vehicle per simulation. Due to the overpredicted settlements the deformation differences proceed faster than in the prototype, so that the calculated behaviour can be considered as long-term behaviour.

A Parameter Study. The calculations have been intended to show the dependence of the settlements of the ballast bed on the initial state. Before the installation of rail and sleepers, the initial state of the track is defined by the initial state of the ballast. In this state the initial ballast height or the initial ballast density have been varied, i.e.

$$h_i = h_0 + \chi_i\,\Delta h;\ i = 1\ldots 60 \qquad\text{or}\qquad e_i = e_0 + \chi_i\,\Delta e;\ i = 1\ldots 60$$
$$e_i = e_0 = const. \qquad\qquad\qquad\qquad h_i = h_0 = const.$$

with h_i the ballast height and e_i the void ration in the ith ballast element, the maximum variation Δh and Δe of height and void ratio resp. and a sequence χ_i of random numbers, $\chi_i \in [-1,1]$. In all simulation calculations in this parameter study the train velocity has been 250 km/h.

a) $\Delta h_{ini} = 0.2$ mm   b) $\Delta h_{ini} = 1.0$ mm
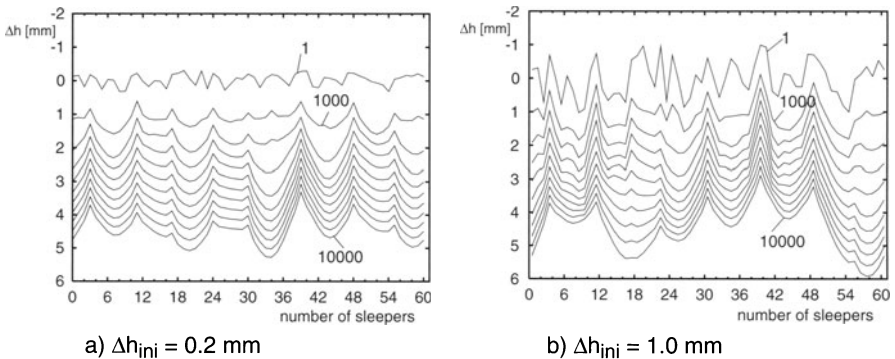
Fig. 22. Ballast height for 1, 1000, 2000, ... 10000 vehicle passages and initial variation of height

Figure 22 shows the ballast height over the number of sleepers for every 1000 vehicle passages in two different cases. In the initial state the ballast height varied along the sleepers and the void ratio was kept constant. We see, that beginning from a random height distribution without any systematic pattern, track bed waves of characteristic length evolve. This length is about 8 to 9 sleeper distances. When a trough has been established it does never transform into a saddle, which is due to positive feedback. As can be expected, an initially stronger variation leads to an evolution of larger height differences. If we compare directly the final states of the two cases, Fig. 23, we conclude that the mean height change is nearly equal, only the differences of height are more extreme in the case of higher initial height variance (dotted line).
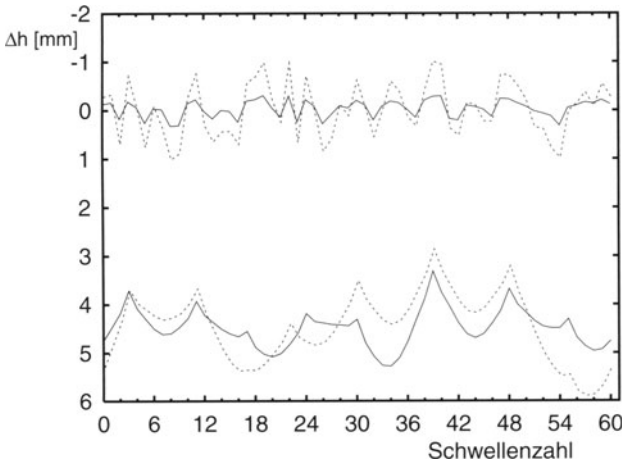
Fig. 23. Initial and final states for different initial height variations

The variation of void ratio e over the ballast elements results into a local variation of ballast stiffness (This kind of variation can be due to tamping). Figure 24 shows the results of two numerical simulations with different initial variations of e.

We see that, as for height variations, after ca. 2000 passages a wave pattern has been established which does not change during the following passages any more. Only the height differences increase. The wavelength is slightly larger, about 11 sleepers distances. The influence of the initial variance of the void ratio on the differences of accumulated deformation is considerable, see Fig. 25. Also the medium height difference in the case $\Delta e = 0.05$ (dotted line) is slightly larger.

A comparison of the two types of initial variation shows that an initial variation of void ratio produces both more mean and more differential height difference. Thus in the process of track construction particular attention should be paid to a uniform high density. In addition we remark that initial differ-
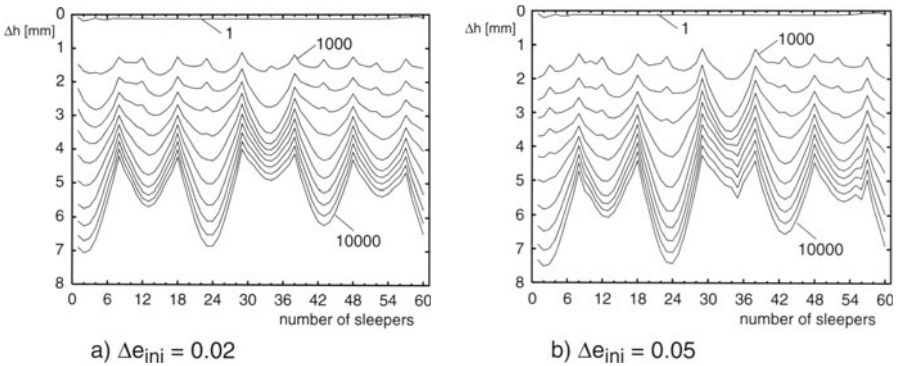
Fig. 24. Ballast height for 1, 1000, 2000, ... 10000 vehicle passages and initial variation of void ratio
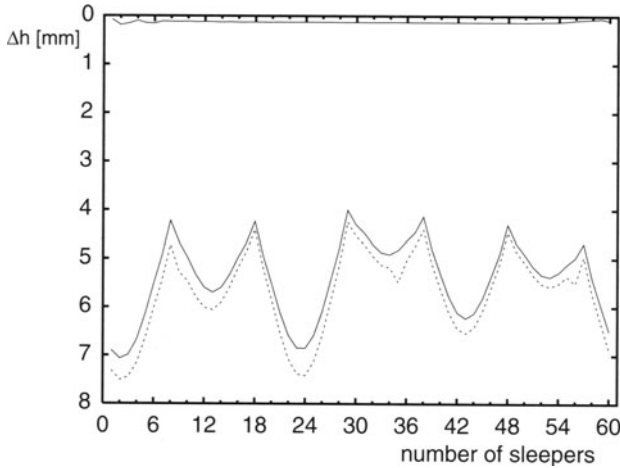
Fig. 25. Initial and final states for different initial variations of void ratio

ences in void ratio result in more pronounced differential settlements than initial differences in ballast height. The simulation results suggest to apply a repair method that does not disturb the stabilized grain skeleton by an alternating dynamic excitation and which avoids a density variation.

Practical experiences show that the evolution of unevenness in the track bed is closely connected with the velocity of the trains. Clearly the acceleration forces play a decisive role in the process. The numerical simulations show that by reducing the velocity of the vehicle the mean as well as the differential settlement of the track bed reduces considerably, see Fig. 26.

Clearly the reasons of this effect are smaller dynamic loads due to the lower speed of 120 km/h.
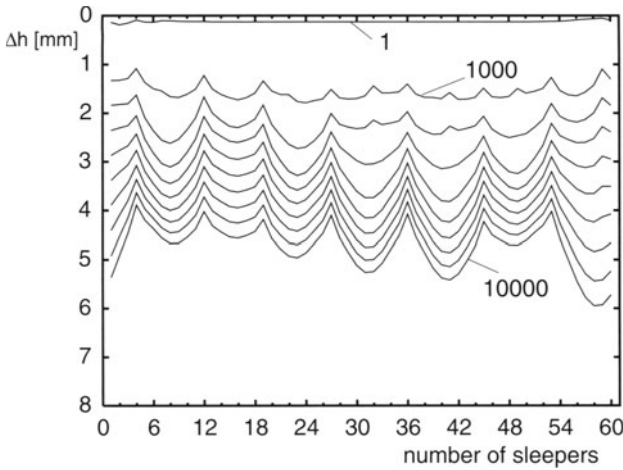
Fig. 26. Ballast height for 1, 1000, 2000, ... 10000 vehicle passages and vehicle velocity of 120 km/h

Summary of the Numerical Track-Train Simulation Results. A numerical model of the track and the vehicle has been worked out in which the ballast material is modelled by a nonlinear constitutive equation, the hypoplastic law. It has been shown that a positive feedback can be found in model tests, element test simulations and in the numerical track model. This explains the growth of height differences in observations and numerical simulations. Some hints for possible repair methods have been proposed. A parameter study has been carried out which shows that a characteristic length of track bed waves evolves. The height of the waves is dependent on the initial variation of the ballast state. Initial density variation results in larger differential settlement than initial ballast height variation.

Acknowledgement

References

1. Augustin, S. (2002) Untersuchungen zur Lagestabilität des Schotteroberbaus. Veröffentlichungen des Institutes für Bodenmechanik und Felsmechanik der Universität Fridericiana in Karlsruhe, Heft 154

2. Bauer, E. (1996) Calibration of a comprehensive hypoplastic model for granular materials. Soils and Foundations, **36**(1), 13-26
3. Froehling, R.D (1999) Measurement and analysis of differential track settlement. In: Schotteroberbau: Mechanische Modellierung, Laborversuche und die Praxis, IFV-Bahntechnik Technische Universität Berlin, Berlin, contrib. VIII
4. Froehling, R.D. (1997) Detoration of Railway track due to Dynamic Vehicle Loading and Spatially Varying Track Stiffness. Ph.D. Thesis, University of Pretoria
5. Gudehus, G. (1996) A comprehensive constitutive equation for granular materials. Soils and Foundations, **36**(1), 1-12
6. Guldenfels R. (1996) Die Alterung von Bahnschotter aus bodenmechanischer Sicht. Veröffentlichungen des Instituts für Geotechnik der EHT Zürich, Band 209
7. Niemunis, A., Herle, I. (1997) Hypoplastic model for cohesionless soils with elastic strain range. Mechanics of Cohesive-Frictional Materials, **2**, 279-299
8. Andreas Schünemann (2002) Numerische Modelle zur Beschreibung des Langzeitverhaltens von Eisenbahnschotter unter alternierender Beanspruchung. Veröffentlichungen des Institutes für Bodenmechanik und Felsmechanik der Universität Fridericiana in Karlsruhe, (in preparation)

Track Settlement Due to Cyclic Loading with Low Minimum Pressure and Vibrations

Matthias Baessler and Werner Ruecker

Federal Institute for Materials Research and Testing (BAM),
Unter den Eichen 87, 12205 Berlin, Germany

Abstract. The paper deals with two specific questions concerning the long-term behaviour of ballasted tracks. We were able to show that the influence of the lower load during cyclic loading is essential for the understanding of track settlement. This is especially true when a gap between sleeper and ballast surface exists. The second part concerns the dynamic excitation of the track. Several excitation cases are discussed and a realistic test sequence is introduced. Test results are described for the following cases: a base excitation of the track, a dynamic excitation of a preloaded sleeper and impacts of a voided sleeper on the ballast. The deformation mechanisms involved in these load types are qualitatively different.

1 Introduction

Ballasted tracks are still the most common railroad track structure. Functions and advantages of ballasted tracks are the distribution of the load from the sleepers, damping of dynamic loads, providing lateral resistance and water drainage, low costs of purchase and the possibility of maintenance operations. Nevertheless, life cycle costs are high and most experts agree that the track maintenance costs have risen with the introduction of high speed trains because of increased loading of and requirements to the track. Therefore, it is important to understand the settlement mechanisms (and the problem which parameters have changed with high speed traffic) in order to obtain improved structures. The usual suggestion that the track settlement depends mostly on the maximum load, the track stiffness and the number of load cycles might not be sufficient and has to be further investigated.

During the last two years of the priority program, we did experimental research at the BAM concerning the settlement behaviour of ballasted tracks. Our aim was to understand fundamental mechanisms of the track settlement. The main parameters we are dealing with are the loading conditions of the track structure applied by the moving trains and later on to some extend different track structures themselves. However, we do not deny that for example the type of ballast material, the grading and the preconditioning and maintenance operations are important parameters as well. A summary of experimental results, railroad settlement models and useful remarks concerning track settlement are given in [1] and [2]. In this research project, we started to build up the necessary database. Furthermore, in the short term we wanted to

clarify the importance of two track loading conditions: the minimum vertical pressure between sleeper and ballast during cyclic loading and the influence of temporary vibrations/impacts on the settlement.

2   Loading Conditions in the Track

2.1   Minimum Load During Cyclic Loading

Under a moving uniform load, the loading of the ballast layer under the sleepers can be seen as nearly cyclic. Also, in reality the profile of the ballast surface is not even. Therefore, when comparing adjacent sleepers one finds that there is a variation in maximum and minimum load. Under voided sleepers, the vertical minimum pressure is zero and in the unloaded condition, there is a gap between the sleeper and the ballast (Fig. 1).
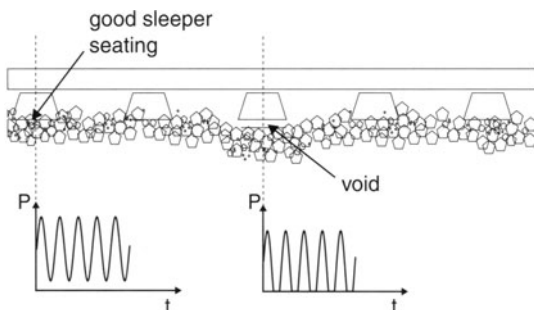
Fig. 1. Loading of sleepers with different seating

For specific axle distances and maybe especially with fast trains, a lifting of the sleepers in front of and behind the wheel is possible. Also, in this situation the minimum load in a load cycle is reduced, at least for a short time (Fig. 4). Concerning these conditions, we wanted to find out if the degree of unloading is a decisive factor and if maybe especially the lifting of the sleeper from the ballast surface is very severe. A first indication that the minimum load is important is given in [3].

2.2   Vibrations and Impacts

The second specific question concerns the influence of vibrations and impacts on the track settlement. As we can see from measurements (Fig. 2), along with the moving static load, the passing of a train is combined with a vibratory loading of the track; this happens because of irregularities of the track or of the wheel. Concerning accelerations, the frequency content can be very high, even above 100 Hz. Also, the moving load means a dynamic loading of the track itself.
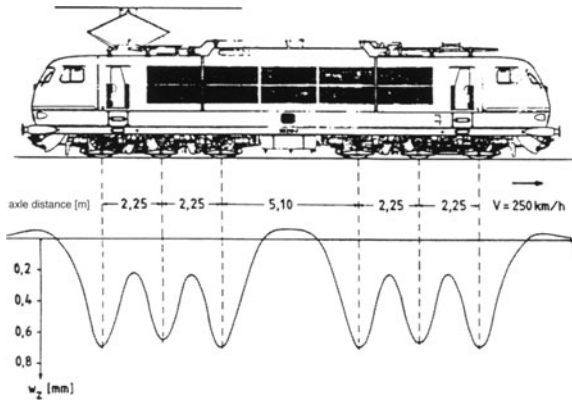
Fig. 2. Deflection curve of an E103 engine
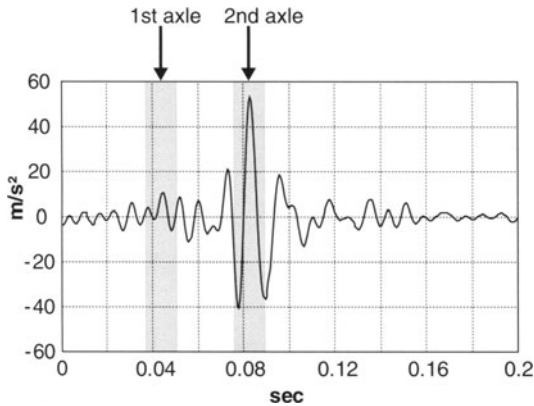
Fig. 3. Acceleration signal from a measurement [4] on a sleeper for an ICE at about 200 kph. Low pass filtered at 200 Hz; acceleration close to 6 g; vibration significant about 80 Hz

Dynamic action, of course, could lead to higher maximum stresses in the ballast layer. Fig. 3 shows additional deformation mechanisms. It is not unlikely that because of the dynamic loads the ballast is pushed aside. A high acceleration level could relieve the slip between the particles or even lead to a jumping of ballast particles. Because of dynamic loading of the track, other mechanisms of permanent track deformation are generally possible, compared to those which could be observed during low frequency cyclic model tests. The increase of vibrations with increased speed of trains is shown in [5]. Findings by other authors concerning dynamic loading will be referred to where appropriate.
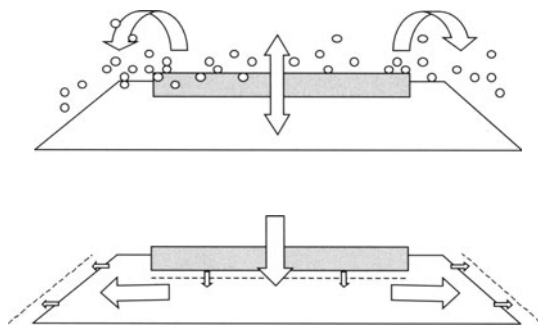
Fig. 4. Principal additional deformation mechanisms because of a dynamic loading of the track; (a) jumping of particles at high acceleration level, (b) ballast pushed aside

3   Experimental Research and Findings

For the experimental investigations two test rigs were used: a relatively large one with a 1:1–model of a four sleeper track segment and simplified tests with a single half sleeper. The test results are presented in chronological order. Therefore, in 3.1 the tests with the large set-up are described. These are the tests we started with and where we learned a lot about the general settlement behaviour. Here, some of the main objectives of our research work are introduced, but descriptions and explanations are kept to a minimum needed for a basic understanding.

We found out that for investigating the specific questions appropriately, the smaller set-up was especially useful. While with the large set-up it was unavoidable to include several loading mechanisms at the same time, the simplified tests provided us with the opportunity to investigate the different load mechanisms separately. Therefore, the extended description of our research and findings is to be found in 3.2 (for the lower load during cyclic loading) and 3.3 (for the dynamic excitation) with the description of the simplified tests.

3.1   First Tests with a Large Track System

Set-up and Test Sequence We started our experiments with a relatively large set-up, which can be seen in Fig. 5. The test rig consisted of a concrete trough filled with soil, tamped in layers by a tamping plate. Then, a ballasted track with four sleepers was built up. The rail was cut in the middle to avoid an unwanted voiding under the immobile load actuator.

The measurement device includes sensors for displacement, velocity and acceleration on the sleepers, pressure cells (by means of strain measurements between stiff plates with a diameter of 10 cm) and accelerometers in the ballast and protection layer, as well as lateral displacements at the ballast
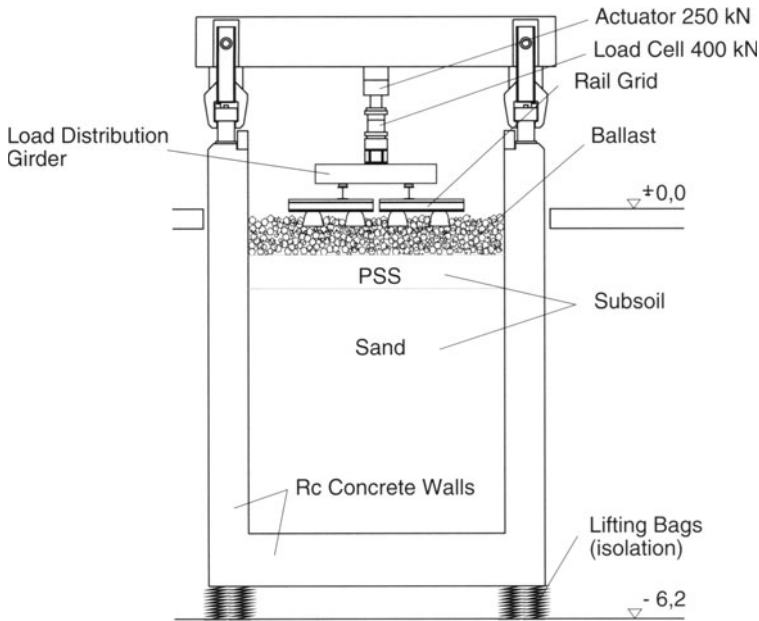
Fig. 5. Assembly of the geotechnical test rig

slope. For the dynamic measurements, the sample rate was usually 2 kHz. Nevertheless, signals in the figures of this text are low-pass filtered at 200 Hz, which is only important for the acceleration measurements. The reason for doing this was to reduce very short peaks which should be meaningless for the settlement behaviour. Also, it should be mentioned that because of the definition of the sign of the load, the most negative value of the load is referred to as the maximum load, the most positive as the minimum load.

For a basic understanding of the settlement behavior of the track, we started with low frequency (1.5 Hz) cyclic loading tests. These tests mean a preloading of the track and the occurring long-term settlement curves are used as a comparison for the very different tests taking place immediately after the cyclic loading. In this next part of the tests we applied certain dynamic load sequences we thought to be realistic.

As outlined before, in reality though, we see short impacts by the sleepers in a certain distance to each other. Dynamic loads are accompanied or followed by nearly static loads. The ballast is generally compacted and constrained. We tried to take these facts into account in our test concept. Therefore, here and in all other tests including dynamic loads we worked with a test sequence as is shown in Fig. 6, with an alternation between dynamic and cyclic loading. Through the cyclic loading, it is guaranteed that the maximum loads of the preloading in part 1 is reached from time to time. The repeated cycles of cyclic loading also give a good basis for measuring settlements and compar-
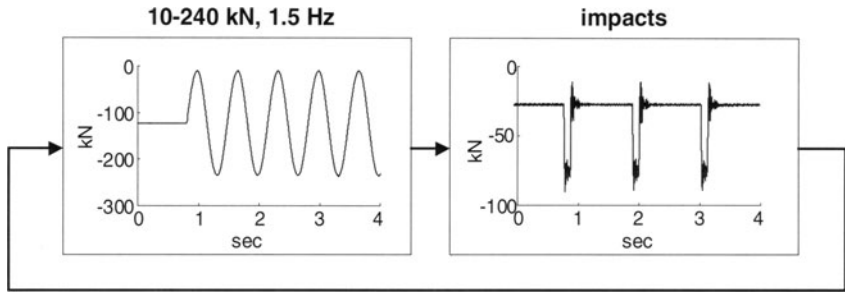
Fig. 6. Example of a test sequence; an alternation between impacts and cyclic loading

ing deformation behavior at several test stages. Following a recommendation in [6], the ballast is preconditioned with a tamping plate, since with sand settlements are highly overestimated for loose fillings compared to preloaded ones (even if the void ratio of the sample is the same).
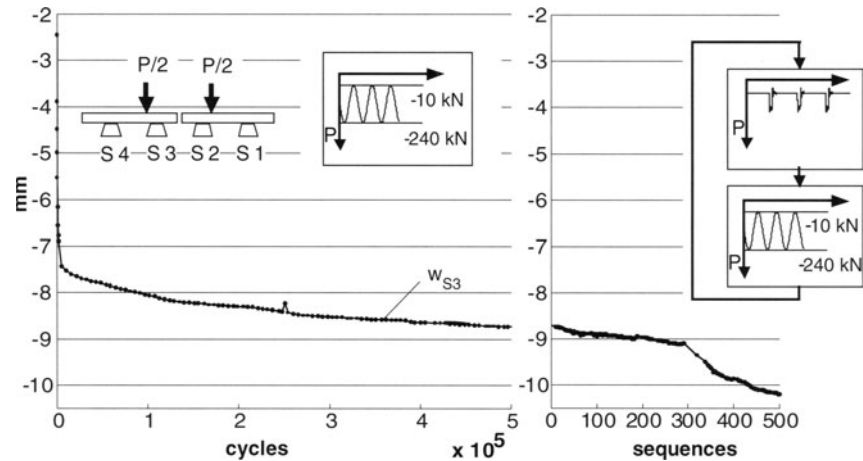
Fig. 7. Settlement for a sleeper (S3) with cyclic and then with an alternation of cyclic and dynamic loading

Test Results An example for the tests we have done is given in Fig. 7. A relatively small increase in permanent deformations is observed after the first cycles are over. Here it is 500.000 cycles under a cyclic load between -10 and -240 kN at 1.5 Hz. The settlement curve would be fitted well by a logarithmic function. The figure on the right shows the settlements under the effect of the impacts mentioned before. We alternated between some impacts and some cycles of cyclic loading. The cyclic load amplitudes are the same we

applied during the first 500.000 cycles. On the horizontal axes the number of such sequences is drawn. 500 sequences only mean some thousand additional cycles. A greater increase in deformations than before can be seen. This must be caused by the impacts.

Having a closer look at details, we did find that during the impacts the maximum load is far smaller than during cyclic loading. This was also true for the stresses we have measured in the ballast layer in lateral and vertical directions. The maximum vertical downwards displacement of sleepers was smaller during impacts as well. But the sleepers are lifted a little more during the impacts than during cyclic loading. Also, high accelerations in the ballast are observed. It was most interesting to see that our initial thesis could be confirmed at a realistic model. The settlements increased whenever dynamic loads were applied and the minimum load was lowered temporarily, although the load, pressure and deflection maxima were not higher than before. Nevertheless, in order to be able to say which one of the two parameters is how important for the settlements, we decided to choose a simplified set-up to investigate the parameters separately.

3.2   Simplified Tests - Minimum Load During Cyclic Loading
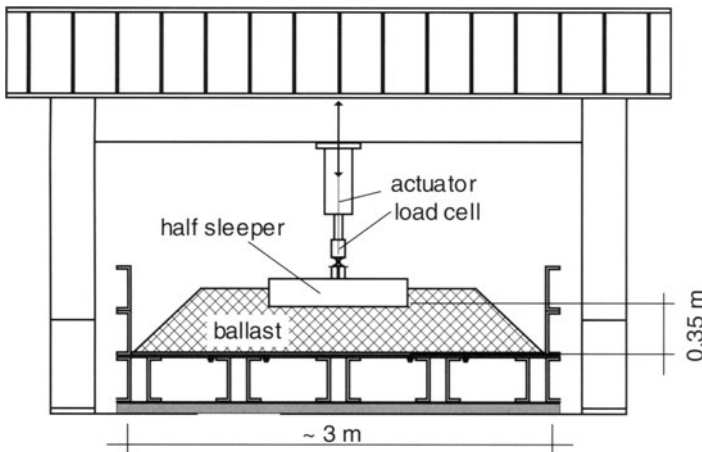
Fig. 8. Set-up for tests with a single half sleeper

Set-up and Test Sequence Fig. 8 shows the second test rig. The loading of the 1:1–track model is brought about by a single half sleeper. For the support of the ballast layer, different stiffnesses are introduced by ballast mats or (as a more rigid support) a plate of stiff wood. A stiff and a joint connection between actuator and sleeper is possible. The load cell is between the sleeper and the load actuator. Therefore, a load of zero indicates that the weight of the sleeper is still on the ballast layer.

The first aim in these tests was to find out more about the influence of the lower load during cyclic loading on the settlement. In several tests we varied the lower load up to a gap between ballast and sleeper as shown qualitatively in Fig. 9. To enable the lifting, test control has to be changed to displacement control with a load limit. With the software used the gap can be kept constant, even when settlements occur.
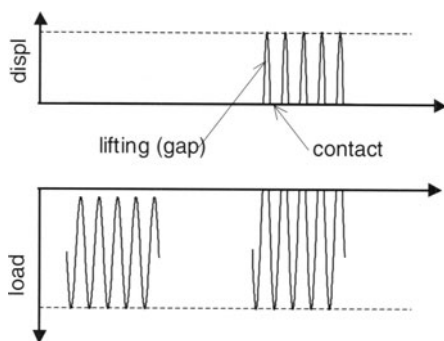
Fig. 9. Load sequence for the tests. Load and displacement between sleeper and ballast. While the maximum load is the same, the minimum load is reduced stepwise until a lifting of the sleeper from the ballast occurs

Test Results The tests carried out clearly show the evidence of the unloading condition. The test shown in Fig. 11 summarizes results obtained in the other tests (support: a plate of hard wood). We started with a low frequency loading from -2 to -30 kN. When we increased the upper load from -30 to -40 kN, deformations also increased significantly, as is well known from other tests. Decreasing the lower load only from -2 to $+1$ kN - which means a near neutralization of the sleeper weight - leads to an increase in settlement again. In the next step, we created a void during unloading by lifting the sleeper from the ballast surface. This led to the largest increase in settlements by far. We also measured lateral deformations at the slope, as can be seen in the third graph of Fig. 11. Horizontal deformations increase, especially when the sleeper is lifted from the ballast surface. The explanation for the large influence of unloading on the settlement can be found in the non-linear load-displacement relation (Fig. 12). A small change in minimum load effects a relatively large change in deformations in one cycle. Rotations and rearrangements of the ballast particles are likely to happen, as we were able to observe in a first test without an embedding of the sleeper, where we had a look at the ballast surface between sleeper and ballast.

The ballast particles certainly move most easily if there is a low minimum load. This also means that at that point any kind of additional disturbance - like horizontal sleeper movements or vibrations - can be extremely severe
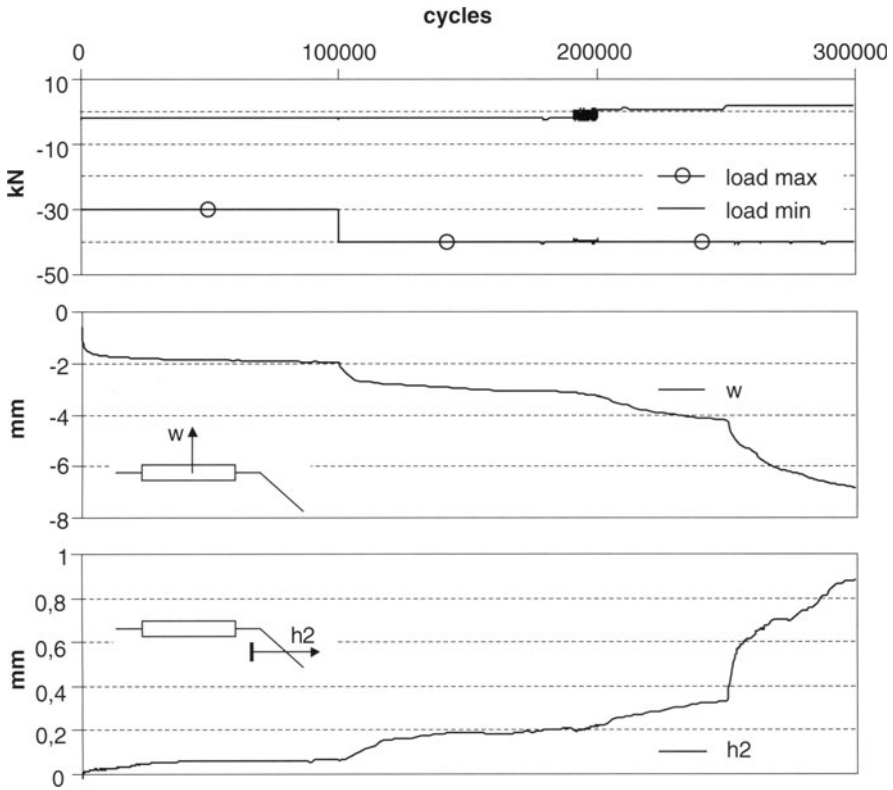
Fig. 10. Permanent deformations due to low load level during cyclic loading; at 250.000 cycles, a gap exists during unloading
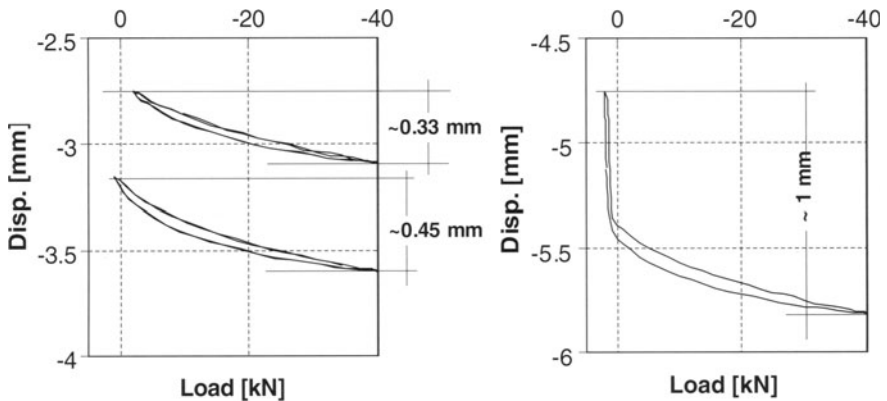
Fig. 11. Load-displacement relation at different test stages with a different load minimum. The figure on the right obviously shows the situation with a gap

concerning the settlement. For instance in [7], it is shown from measurements that especially the lateral dynamic motion of the sleepers increases with the track irregularities. Also, because of temperature changes and loading the rail deforms laterally, for example at bridges, as is reported in [12]. This will also affect the sleepers and thereby the ballast surface. Again, particles could be rearranged, leading to increased settlements.
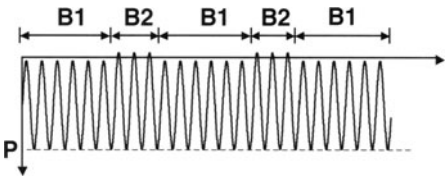
Fig. 12. Change of lower load during cyclic loading

Table 1. Load sequence for temporary track lifting

series	$p_{MIN}[kN]$	$p_{MAX}[kN]$	cycles
A	−2	−40	-
B1	−2	−40	200
B2	+1	−40	5
C	+1	−40	-

Temporary Track Lifting Even in a perfect track, a lifting of a sleeper can happen temporarily. Whether this is important in any way can be seen when looking at the sequence in Fig. 13. Firstly, we had the cyclic loading from −2 to −40 kN. Then we alternated between more and less unloading. The sequence can be seen qualitatively in Fig. 12. Following Tab. 1 it is 200 cycles between −2 and −40 kN (B1) and then 5 cycles +1 to −40 kN (B2) and so on. We found that even the temporary appearance of some sequences with a lower load lead to an increase in settlements.

3.3   Simplified Tests – Dynamic Excitation

The other topic we dealt with is the influence of dynamic excitation on the track. We did several experiments with base excitation, dynamic excitation brought about by a sleeper or, as a special case, with impacts of a voided sleeper.
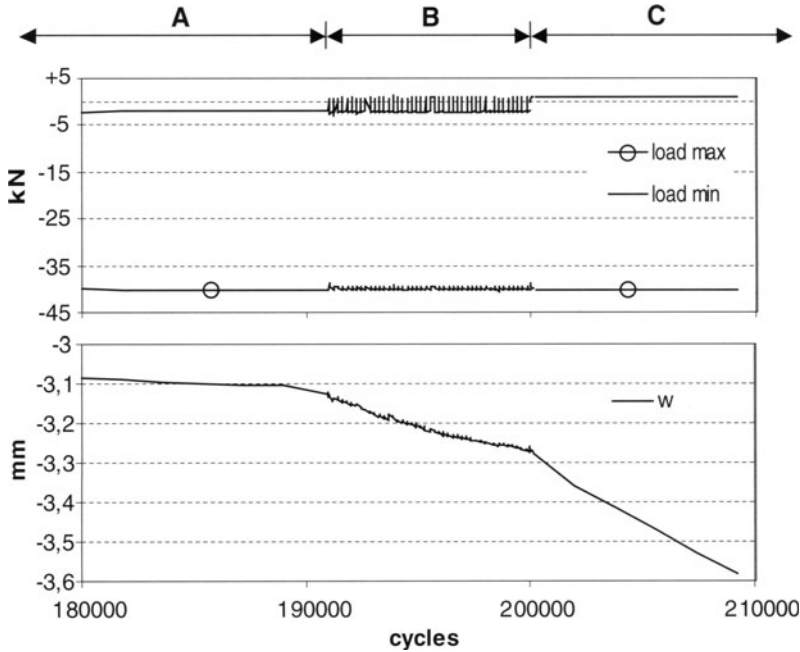
Fig. 13. Load amplitudes and maximum sleeper displacements for a temporary track lifting

(a)   (b)   (c)
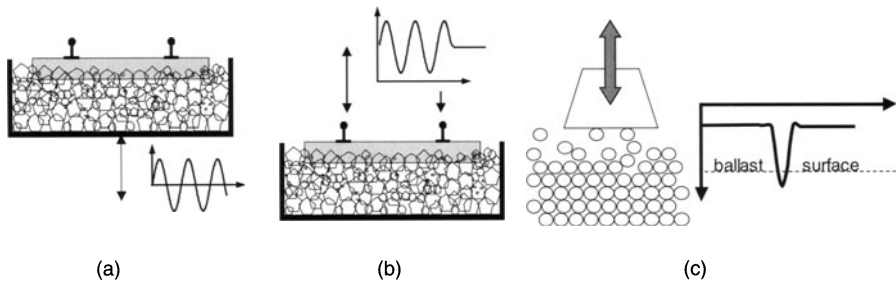
Fig. 14. (a) Base Excitation, (b) excitation from the top, (c) impacts of a voided sleeper

Introductory Remarks Some research has been done on the influence of dynamic loading of the track. Firstly, the increase in maximum pressure because of the increased dynamic loads is taken into account. This can be done with nearly static cyclic model tests. Therefore, for instance [8] finds an increase in settlements with the power of 1.6 of the maximum load (conclusion in [1]). Other researchers do tests with a harmonic base excitation of a ballast sample. Generally, the result is a threshold value for destabilization and settlements of about 1 g ([9], [10] and literature cited there). This is caused

by the fact that in this case the ballast cannot follow the excited table in the downwards movement.

Base excitation of a track is sometimes to be found on short bridges for high train speeds when a resonance of the bridge deck follows the train passing. In our own tests with base excitation of a ballast sample [11] we also found a destabilization of the ballast at least at about 1 g, but earlier for higher frequencies (frequencies up to 50 Hz). The destabilization includes a rapid change in particle contacts which sometimes is referred to as fluidization. For higher frequencies, there is an amplification of the vibrations from the table to the top of the track. The tests were carried out with a 1:1–model of a bridge segment with a four sleeper rail grid. For a track model with a slope we observed a flattening of the slope for higher accelerations. Obviously, the particles on the surface slide and roll down (lateral flow). Another impression is that higher frequencies are not so severe at the same acceleration level in the ballast layer, but this is hard to prove by measurement data with these tests. More research is needed to clarify this point. Generally, the observed destabilization does not say very much about the amount of settlements, which could occur in reality due to this excitation. In reality the track loading on bridges is an alternation and interaction between the base excitation of the bridge desk and the cyclic loading of the track from the top, which is missing here. Also, the tests are done with a compacted, but not preloaded ballast sample. Nevertheless, of course an excitation that could lead to destabilization – the point we were investigating – has to be avoided.

Generally, our main objective here is an excitation from the top. This is certainly more difficult to understand and tests are more difficult to perform. At this point it is only possible to show general trends. From a certain excitation level particles on the top will lose contact temporarily. If so, then the motion of the particles is not periodic anymore. The flight height can increase with the number of load cycles as is shown for example in [13]. As said before, in measurements we see a variety of short impacts and more or less cyclic loading. If the dynamic action loosens the ballast, then the cyclic loading leads to an increase in confining pressure again. This has to be taken into account. Furthermore, in a test box, standing waves can occur due to the harmonic excitation. Therefore, our conclusion is that for the simulation of other dynamic loading conditions in the track than that on bridges, *a permanent harmonic excitation of the track model is usually misleading*. Also, in contrast to the situation on bridges it is not the whole sample that is shaken uniformly, but different parts in the layer behave differently in magnitude and phase.

Another problem is also the fact that it is difficult to separate the different mechanisms leading to increased settlement. The influence of the increased maximum load on the settlement is well known. If you want to know if the additional mechanisms of pushing aside or the jumping of ballast particles is important, you have to avoid load conditions where just the level of ver-

tical stress or strain is higher than before. Due to inertia this implies not only a look at the load measured above the sleeper, but also on the pressure measurements at the bottom of the ballast layer. In the test sequence used mechanisms are separated. The maximum load exists during the cyclic loading while additional mechanisms are involved for the dynamic loading.

Apart from the base excitation of a track mentioned before, we investigated two types of dynamic excitation from the top: impacts of a voided sleeper and vibrations of a sleeper with a good seating.

Impacts of a Voided Sleeper A first example for the dynamic loading are impacts brought about by a voided sleeper (Fig. 14 c). The influence of lifting during cyclic loading has been shown before. But for some cases voided sleepers are sometimes pushed onto the ballast layer surface. Supplementary to the lifting, results concerning this loading are shown here.
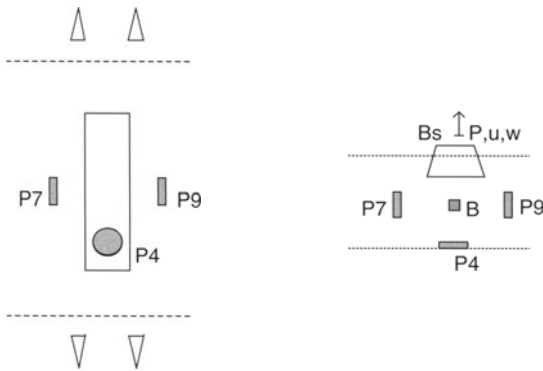
Fig. 15. Top view and cross-section of sensors referred to in the text

In the test sequence we used, we pushed the voided sleeper 5 times on the ballast surface. This is followed by a cyclic loading (0.5 Hz) between lifting (voided sleeper) and -40 kN. Positions of the sensors referred to can be seen in Fig. 17. Displacement and loading amplitudes (Fig. 15 a,b) are far smaller for the impacts than during the nearly static loading. But accelerations measured in the ballast layer are high, even if only short peaks occur (15 c,d). The question was: What happens as a consequence of the impacts? In the tests we started with some thousand cycles of cyclic loading with the amplitudes mentioned before. Then impact sequences are followed by the same cyclic loading (30 cycles). These sequences are repeated about 50 times. The settlement increases whenever we apply the impact loading (Fig. 18).

Some details give a little insight into the mechanisms. Fig. 16 shows a part of the loading sequence. The cyclic loading is followed by some impacts which are again followed by cyclic loading. The lower figure shows the load displacement relation at several stages of the test, marked differently. What we
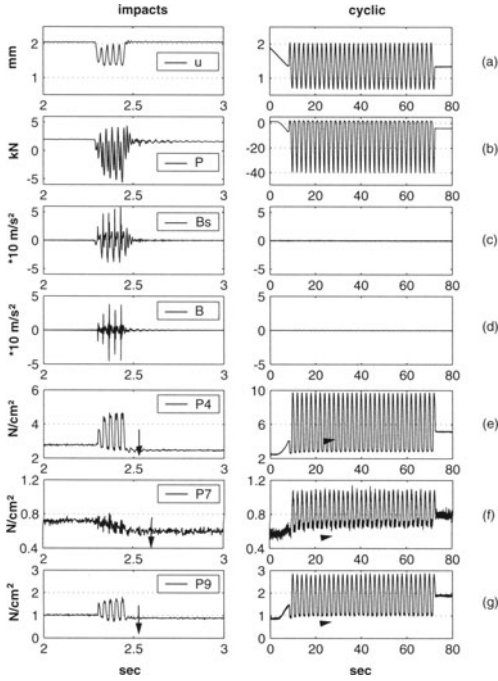
Fig. 16. Time signals for impact and cyclic loading of the track. Displacement load actuator (a), load (b), accelerations on sleeper (c) and under sleeper (d), pressure measurements vertical (e) and horizontal (f, g) in the ballast layer

can see are nearly equal curves for the final cycles of each sequence of cyclic loading. A little settlement increment can be guessed at. The dotted curve follows directly after the impacts. Here, the level of the ballast layer has lifted and the stiffness at loading is obviously reduced. The ballast has loosened a little because of the impacts. The latter fact can also be seen from the pressure measurements we did. During the impacts, the stresses are reduced a little (Fig. 15 e, f, g). During the cyclic loading they increase again. The settlement does not increase immediately after the impacts, but needs some cycles of cyclic loading. Due to the impacts the rearrangement of particles is relieved.

Dynamic Excitation from the Top We also did some experiments with a dynamic loading from the top and a good sleeper seating. After 25.000 cycles
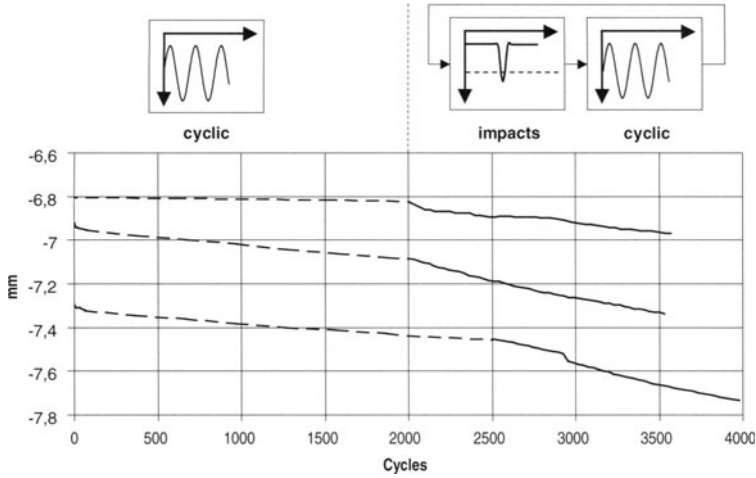
Fig. 17. Increased settlement after sequences with impacts are applied. Only cycles of cyclic loading are counted
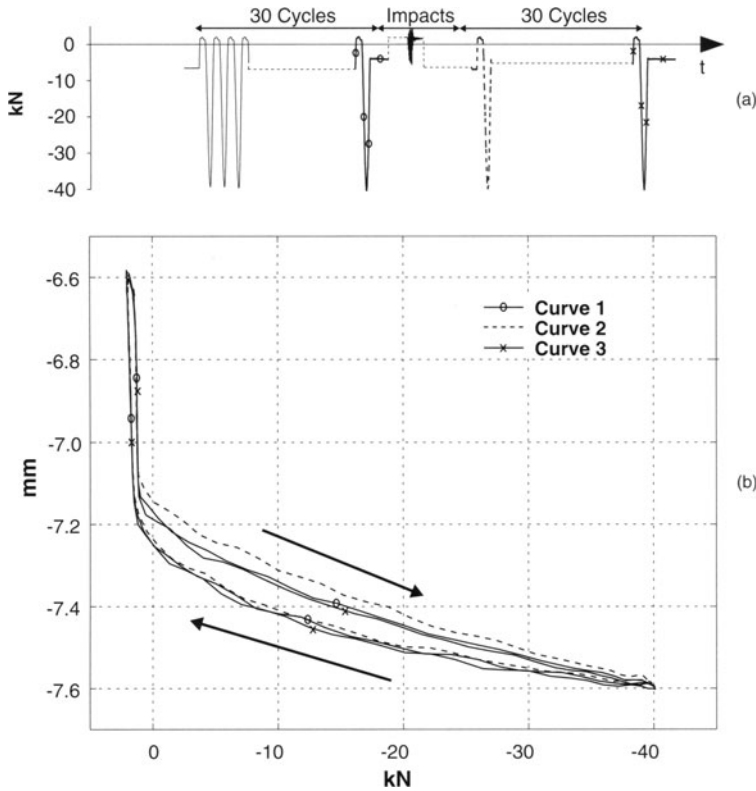
Fig. 18. Load–displacement relation for different test stages (b) corresponding to the sequence (a)

of cyclic loading for comparative use, we introduced sequences with low load amplitudes at higher frequencies. We varied the frequency, the amplitude and the number of cycles to get a first impression about key parameters influencing the settlement behaviour. Sequences with dynamic loads are again alternated with 30 cycles of cyclic loading. The cyclic loading for the preloading and after each dynamic event has again an amplitude between −40 kN and a lifting of the sleeper from the ballast surface. This is done to guarantee that increased settlements do not just occur because the sleeper is lifted more during the dynamic loading. In further tests we will change that. For the five test series basic data is given in table 2.

Table 2. Data for the dynamic loading during test series S1 to S5

series	$f[Hz]$	$p_0[kN]$	Δp	cycles
S1	80	10	Fig. 19 a	2
S2	80	10	Fig. 19 b	2
S3	80	10	Fig. 19 c	2
S4	30	10	Fig. 19 d	2
S5	30	10	Fig. 19 e	20

The load amplitudes behind this data can be seen from Fig. 19. The load and an acceleration signal (measurement in ballast layer under the sleeper) are shown to characterize the loading. The intermediate cyclic loading is the same as in Fig. 15. Three test series were performed with 2 cycles at 80 Hz (Fig. 19 a–c). The amplitude is increased for every series. What can be seen from Fig. 20 is a strong increase of settlements for the highest acceleration level, where the peak is about 4 g. The tests do not indicate, that 1 g is a threshold value from which settlements increase rapidly, as is sometimes mentioned. Test series S4 and S5 were done with an excitation at a lower frequency of 30 Hz. The difference in these series was the number of cycles which is changed from 2 to 20. Also, comparing the series, at S4 the amplitude is little higher than during S5. Nevertheless, the increase in settlements is more rapid for the more permanent excitation in the series S5. Even S4 leads to an increased settlement, although the acceleration level is not so high compared to S1–S3. Therefore, a first simplified conclusion could be that settlements naturally increase with the amplitude of excitation and that lower frequencies have a more deteriorating effect.

An approach where settlements increase with accelerations could be useful (as for example in [14], [15]), but it is questionable that the frequency content is unimportant. The deformations for accelerations at high frequencies become so small that this must be a relevant influence. Accelerations can be very high
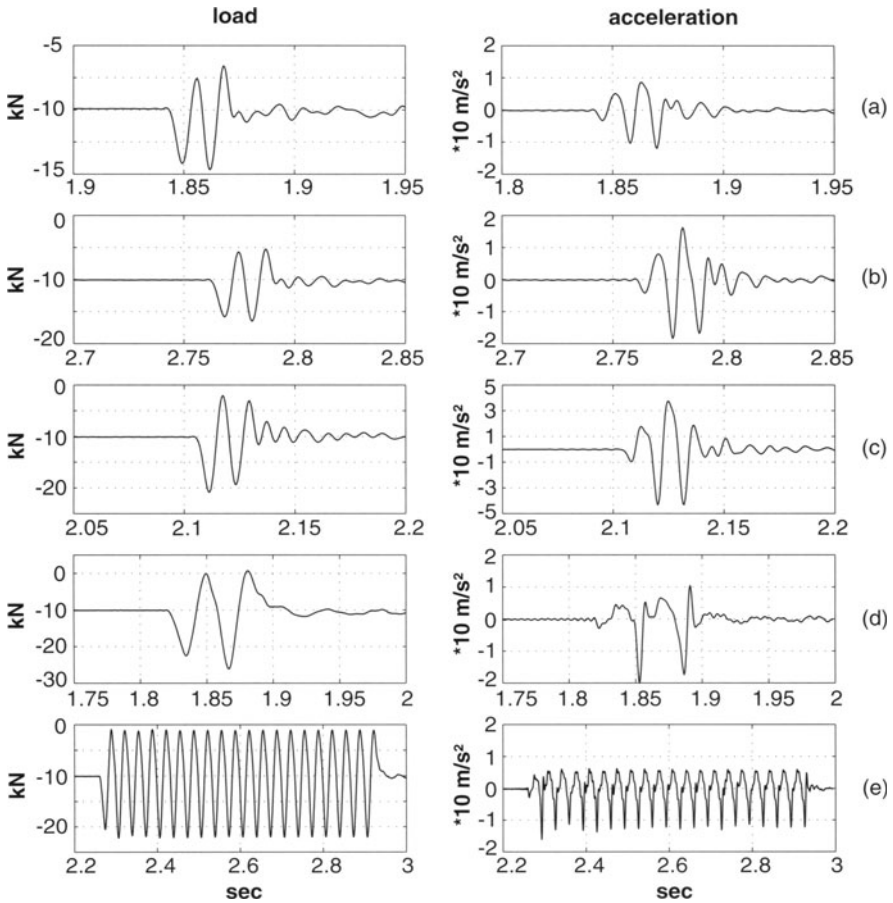
Fig. 19. Load and acceleration measurements corresponding to Fig. 20; (**a**) series S1, (**b**) series S2, (**c**) series S3, (**d**) series S4, (**e**) series S5

in a real track. By the institute of the authors measurements were performed where accelerations as high as 17 g on a sleeper occured. This usually happens because of single peaks and belongs to a frequency content far above 100 Hz. Therefore, the peak is also very much dependent on the sample rate. It is necessary to learn more about how to assess a measured signal with regard to the settlement. For instance in [16], a procedure is described which looks at the energy level of lower frequencies of the vibration close to the track. It could be observed that a lateral flow of particles is most common at places with a high level at these low frequencies and therefore has to be avoided. As said before, a value of about 1 g is not a threshold value. This can also be understood from the fact that acceleration amplitudes reduce rapidly with the distance from the sleeper. Even accelerations in the ballast layer under the sleeper are slightly lower than on the sleeper (as long as there is a permanent
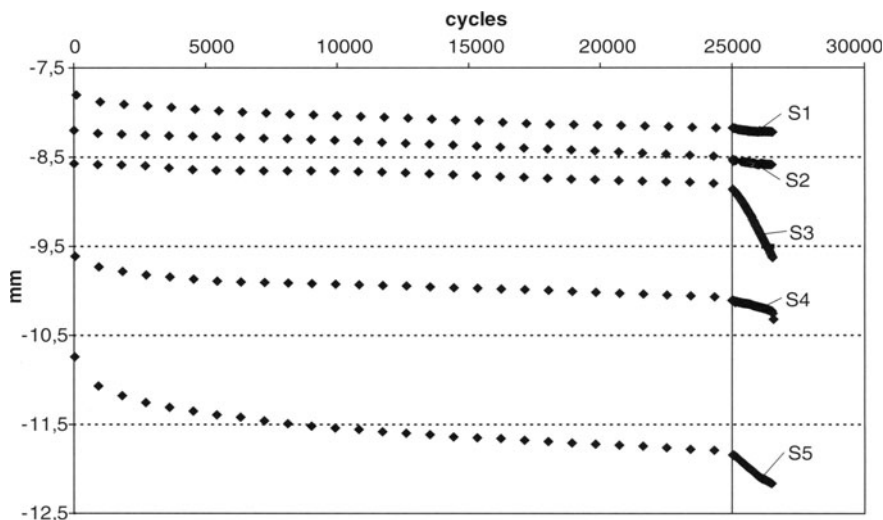
Fig. 20. Settlements for the test series S1–S5; only cyclic loading for 25.000 cycles, then dynamic sequences

contact between sleeper and ballast layer). Where they are highest under the sleeper, the ballast is confined and can not rearrange easily. So it is not likely that the influence of dynamic excitation can be characterized by one or two single values.

4   Summary and Outlook

In our application to this research project two years ago, we formulated the thesis that the ballast layer could react very sensitively to a low minimum vertical load during unloading which again could lead to increased settlements. This text deals with some basic mechanisms we observed with this low minimum stress. What we did see is that settlements are influenced strongly by the level of unloading. A reduction of the minimum load usually leads to a larger increase in settlements than an equal-sized increase in the maximum load does. Under voided sleepers, ballast might not receive the maximum pressure, but deformations are more likely to occur. This is also true if a lower degree of unloading is only included from time to time. The clarification of the influence of the lower load level during cyclic loading on the settlement process has strong implications on the understanding of the long-term behaviour of the real track.

Dynamic effects increase the settlement process of the track. Examples for dynamic loading are the impacts under a voided sleeper, the base excitation on some bridges or just the moving vibratory load of the train. We did find that a dynamic excitation can not be simply modelled by the use of higher

maximum loads. Concerning the impacts of a voided sleeper we were able to show that the ballast particles are loosened and therefore the rearrangements are relieved. Because of the impacts the settlement is further enlarged at voided sleepers (compared to cyclic loading). First test with a dynamic loading for a good sleeper seating indicate that increased settlements occur because of the dynamic action, even if the force applied is small. Taking the acceleration level for comparative use, one finds that higher amplitudes and lower frequencies and the number of dynamic cycles are decisive parameters. However, these are first results which have to be further validated and complemented.

In general, settlement equations only including the maximum load and a track stiffness parameter are lacking some main aspects of track settlement. Improved settlement laws and suitable numerical procedures are needed. For the cyclic loading of the track, advanced continuum descriptions are developed in Karlsruhe and Kassel (see their reports in this volume). If the vibration level increases, it can become necessary to take into account the single ballast particles. For instance in Hannover (also this volume), a two-dimensional model with discrete elements has been used to investigate track settlement. While there is certainly still a lot of development and validation necessary, this might be the best option for a numerical investigation of the dynamically loaded ballast layer. Validation of a discrete particle model at simplified tests is a challenge for future research work.

Especially concerning dynamic loading, more research still needs to be done to clarify which conditions lead to increased settlement. This is important to obtain improved track structures in the future.

References

1. Dahlberg, T. (2000) Some Railroad Settlement Laws–A Critical Review. Handout Distributed by the Author, Euromech Colloqium 409, Hanover, March 6–9
2. Holtzendorff, K. (1999) Ursachen von Schottersetzung und deren Modellierung. In: Schotteroberbau – Mechanische Modellierung, Laborversuche und die Praxis, IFV Symposium, Berlin, February 18–19, 1999.
3. Jeffs, T., Marich, S. (1987) Ballast Characteristics in the Laboratory. Railway Engineering Conference, Perth, September 14–16, 1987
4. Auersch, L., Said, S., Ruecker, W. (2001) Das Fahrzeug-Fahrweg-Verhalten und die Umgebungserschuetterungen bei Eisenbahnen, Verlag fuer neue Wissenschaften, BAM–Forschungsbericht **243**, Bremerhaven

5. Pruefamt fuer Grundbau, Bodenmechanik und Felsmechanik der TU Muenchen (1989) Fahrzeug-Komponenten-Erprobung fuer HGV mit 250 km/h, Boden-mechanische Messungen bei Fahrten mit ICE und Lokzug in den Jahren 1987 und 1988

6. Holzloehner, U. (1997) Schaedigungsmechanismen, Abtragung von Verfor-mungen und Belastungen. Unpublished Manuscript of Paper Held at the Re-search and Development Conference of the Deutsche Bahn, Munich, July 3–4

7. Rump, R., Ehling, B., Rehfeld, E. (1996) Wirkungen von Verkehrserschuet-terungen auf Erdbauwerke und ungebundene Tragschichten im Oberbau. ETR Eisenbahntechnische Rundschau 45, 485–491

8. Hettler, A. (1986) Bleibende Setzungen des Schotteroberbaus. ETR Eisenbahn-technische Rundschau 33, 847–854

9. Gaskin, P., Raymond G., Powell, A. (1978) Response of Railroad Ballast to Vertical Vibration. Transportation Engineering Journal of ASCE, TE 1, 75–87

10. Morgan, J. G. D., Markland, E. (1981) The Effect of Vibration on Ballast Beds. Geotechnique 31, No. 3, 367–386

11. Ruecker, W., Rohrmann, R. et al (1998) Investigation of Ballast Behaviour in Bridges Due to High Acceleration on a Test Rig, BAM, Berlin

12. Eisenmann, J., Rump, R. (1997) Ein Schotteroberbau fuer hohe Geschwindig-keiten. ETR Eisenbahntechnische Rundschau 46, 99–108

13. Kutzner, C. (1962) Ueber die Vorgaenge in koernigen Schuettungen bei der Ruettelverdichtung. PhD Thesis, Technische Hochschule Karlsruhe

14. Sato, Y. (1995) Japanese Studies on Deterioriation of Ballasted Track. Vehicle System Dynamics 24, 197–208

15. Holtzendorff, K., Gerstberger, U. (2001) Predicting Settlements of Ballasted Tracks due to Voided Sleepers, WCRR, Koeln, Nov. 25–29

16. Reamer, S. K. (2000) Auswertung von Schwingungsmessungen zur Unter-suchung der dynamischen Stabilitaet des Untergrundes. VDEI Fachtagung Bahnbau, September 12–15, Berlin, 128–132

Simulation of the Dynamic Behavior of Bedding-Foundation-Soil in the Time Domain

Mohammad Firuziaan and Otto von Estorff

Technical University Hamburg-Harburg, Mechanics and Ocean Engineering, Eißendorferstr. 42, 21073 Hamburg, Germany

Abstract. The simulation of the interaction between a railway bed and a train calls for the development of a reliable computational model. Whereas models for the train, developed with conventional methods, exist and are applied successfully in industry, there is no suitable model to describe, in detail, the dynamic behavior of the bedding, the foundation, and the undisturbed soil.

A simple possibility to take into account these subsystems is to assume an elastic embedment (Winkler model). This model, however, does not fully represent the elasticity and damping, especially in the case of wave radiation. Therefore, the computational results obtained with this approach do not correspond well to measurements.

The present project deals with a combination of the boundary element method (BEM) and the finite element method (FEM) for the numerical analyses of bedding-foundation-soil interaction problems. Different material laws, e.g., an elastoplastic and a damage model, are implemented in order to investigate the influence of different nonlinear models on the dynamic interaction. The accuracy and applicability of the model is shown by several examples, namely a benchmark problem, a tunnel system, and a number of railway track configurations.

1 Introduction

The dynamic behavior of a railway track system often is of particular importance. In the construction phase, for instance, its short and long term behavior under dynamic loads have to be considered and it must be examined if the underlying soil needs to be taken into account. Unless there is a massive rock foundation, the soil usually has a significant influence on the railway track behavior.

Up to now, numerous techniques have been developed which allow to take the soil into account. However, these are employed mainly for the investigation of the dynamic interaction between structural systems and the soil, and all have their merits and drawbacks. Thus, the so-called "layer method", as suggested by Kausel and Roësset [16], Lysmer and Kuhlemeyer [25], Lysmer [24], and Waas [37], can be applied only to a horizontally layered unbounded medium. Other methodologies, which use an approximate boundary condition at the edge of a finite element mesh, often do not take into account the properties of the unbounded medium in a proper way. In most cases, these

so-called "transmitting boundaries" are based on the theory of wave propagation, enforcing outgoing plane waves [17]. An interesting approach has been developed recently from the standard layer methods discussed above. It is the so-called "thin layer method", which can be used efficiently also for the modeling of a stratified halfspace. The methodology is described by Kausel [18] and successfully applied, e.g., by Savidis et al. [33] [34].

Further promising approaches, including simple analytical models as well as BEM and FEM approximations, are discussed in detail by Wolf [39] and Wolf and Song [38], [40]. In these references also many additional research activities by other authors are discussed which shall not be repeated here.

Aiming to improve the modeling of railway track systems, several developments have been carried out up to now. As a representative publication of those models, where a track system is introduced without using a continuum formulation, Ripke's dissertation [31] should be mentioned. Based on a model for the dynamic of an infinitely long rail, Ripke developed a train-track model of finite dimensions, where the foundation has been taken into account using special bedding numbers for the ballast.

Some remarks on the importance of a proper modeling of the bedding-foundation-soil system are given by Knothe and Wu in [20]. They compare the vertical dynamic behavior of a track system, which is, on one hand, coupled to a halfspace, on the other hand just modeled as a viscoelastic foundation. Particularly in the frequency range below 250 Hz, significant differences between the two models could be observed and it was shown that the halfspace model yields more accurate results then the viscoelastic one.

For taking into account the bedding, the foundation and the undisturbed soil, the usage of discretization methods, such as the FEM and the BEM, is gaining more and more of importance. The major reason for this trend can be seen in the fact that these methods are rather accurate and provide a high flexibility when complicated subsystems have to be modeled. The drawback of a rather long computation time loses in significance in view of the rapidly growing computational power and the continuous improvement of the computation algorithms (see, e.g., [19]).

A proper track model needs to meet certain requirements. In particular, the behavior of the rail-sleeper system, the ballast, the subgrade, and the soil needs to be represented [32],[36]. Moreover, if the nonlinear behavior of some of these components shall be taken into account, only formulations given directly in the time domain can be used. For subsystems of finite extension, the use of finite elements seems to be advantageous as shown, e.g., by Luo, Yin und Hua [23] in the case of the ballast. Difficult, however, is the modeling of the "infinite" soil with finite elements. Here the usage of infinite elements could be a solution. Such an approach, however, does not exist as a transient formulation yet, and even in the frequency domain the formulation of such elements is still rather limited (see, e.g., [22]).

The modeling of the infinite soil is mostly done by means of boundary ele-

ments [4]. This implies, however, that the soil medium is homogeneous and behaves linearly elastic. Material damping can be taken into account, as suggested by Gaul et al. [13],[14].

In view of these aspects, the idea of coupling FEM and BEM directly in the time domain was born. With such a formulation inhomogeneities and non-linear effects could be taken into account in the finite element submodels whereas boundary elements were employed to represent the infinite soil [10]. Alternatively, Savidis, Bode and Hirschauer [33],[34] suggested the usage of a halfspace consisting of a number of thin layers instead of boundary elements. An important issue with respect to a complete track model is the loading conditions induced by a moving train. Lefeuve-Mesgouez [21] as well as Pflanz, Huber and Schmid [29] investigated in detail the influence of moving loads, i.e., a passing train. Moreover, Adam, Pflanz und Schmid [1] investigated criteria which might be used to decide whether a 2D or a 3D model might be used.

In all of the research activities discussed so far, it is assumed that the FE subsystems behave linearly elastic or viscoelastic.

This contribution deals with the development of a new model for the investigation of complex (nonlinearly behaving) track systems which is based on a coupling of FEM and BEM directly in the time domain. The FEM is used for the discretization of the bedding, the foundation, and a part of the soil in the near field, while the BEM is employed to model the far field of the soil as a semi-infinite subdomain (see Fig. 1). In the finite element model, inhomogeneities, nonlinear material laws and large deformations can be taken into account. The new approach is applied to a number of two- and three-dimensional examples in order to show its accuracy and applicability. It should be mentioned that further details of the formulation and additional results can be found in a number of papers [6],[7],[8],[9],[11], which have been published during the course of this project.
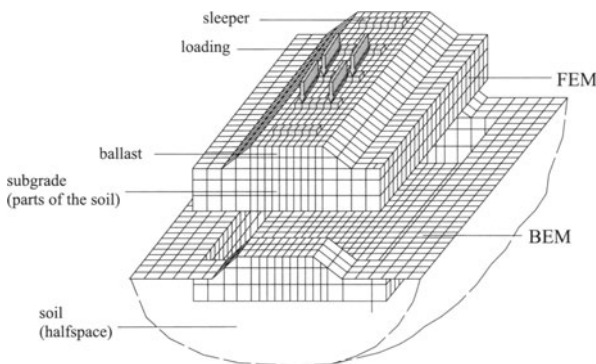
Fig. 1. Discretization of a railway-track system with FEM and BEM

2   Computational Model

In this section the computational model is briefly introduced by summarizing the basic equations of the FEM and the BEM, followed by an explanation of the coupling strategy. To give an idea of possible improvements with respect to computation time, some corresponding aspects are discussed as well.

2.1   Finite Element Method

The starting point for the finite element formulation is the weak form of the dynamic equilibrium [2]. It can be formulated in the undeformed reference configuration Ω_0 of a volumetric body, such that

$$\int_{\Omega_0} S_{ij} F_{ki} w_{k,j} dV + \int_{\Omega_0} \rho_0 \ddot{u}_i w_i dV + G_{ext} = 0, \tag{1}$$

where $F_{ij} = dx_i/dX_j$ represents the deformation gradient (X_j and x_i are the components of the vector defining the location of a point of the body at the beginning and in the current state of the calculation, respectively) and w_i is an admissible test function; ρ_0 is the material density in the undeformed state of the body, while $\ddot{u}_i$ indicates its unknown acceleration field. The virtual work of the external loading is given by G_{ext}. Finally, S_{ij} indicates the second Piola-Kirchhoff stress, given by

$$S_{ij} = \frac{\rho_0}{\rho} F_{ik}^{-1} \tau_{kl} F_{jl}^{-1}, \tag{2}$$

where ρ is the density of the deformed body, $F_{ij}^{-1} = dX_i/dx_j$ is the inverse of the deformation gradient and τ_{ij} are the coefficients of the Cauchy stress tensor.
A linearized form of equation (1) is obtained by using its first variation with respect to the state variable u. Assuming that the load vector does not change its direction throughout the simulation, this linearized form can be written as

$$\delta G = \int_{\Omega_0} \delta S_{ij} F_{ki} w_{k,j} dV + \int_{\Omega_0} S_{ij} \delta F_{ki} w_{k,j} dV$$
$$= \int_{\Omega_0} C_{ijml} \delta E_{ml} F_{ki} w_{k,j} dV + \int_{\Omega_0} S_{ij} \delta F_{ki} w_{k,j} dV, \tag{3}$$

where C_{ijkl} is the material tensor. This can be determined by differentiating the Piola-Kirchhoff stress tensor $\boldsymbol{S}$, given in equation (2), with respect to the Green-Lagrange strain tensor $\boldsymbol{E}$, defined by

$$E_{ij} = \frac{1}{2}(F_{ki} F_{kj} - \delta_{ij}). \tag{4}$$

One obtains

$$C_{ijkl} = \frac{\delta S_{ij}}{\delta E_{kl}}. \tag{5}$$

Note that in equation (3) the variations $\delta F_{ij} = \delta u_{i,j}$ and $\delta E_{ij} = \frac{1}{2}(\delta F_{ki} F_{kj} + F_{ki} \delta F_{kj})$ are used. Moreover, in equation (4) δ_{ij} is the Kronecker symbol, defined by $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$.

The major advantage of the formulation briefly summarized above can be seen in the fact that all integrals of equation (3) can be evaluated with respect to the undeformed configuration at time t_0 (Total Lagrange Formulation). Further details can also be found in [2].

The basic problem in transient nonlinear analyses is to find the state of equilibrium of a body subjected to an external time dependent load. To solve such a problem, we assume that the solution for the discrete time t_i is known and that the solution for the discrete time $t_{i+1} = t_i + \Delta t$ is required, where Δt is a suitably chosen time increment. For each time step, the equilibrium conditions of a nonlinear finite element system representing the body can be expressed as

$$\boldsymbol{R} - \boldsymbol{F} = \boldsymbol{0} \tag{6}$$

where the vector $\boldsymbol{F}$ contains the sum of all externally applied nodal forces. In the vector $\boldsymbol{R}$, on the other hand, the nodal forces, which correspond to the element stresses and the dynamic part of the equilibrium due to inertia effects in the current configuration, can be found.

Using the linearized equation (3), an effective tangent stiffness matrix $\boldsymbol{K}_{eff}$ can be derived which corresponds to the geometric and material conditions at time t and also contains the mass matrix. Since this stiffness matrix depends on the deformation of the system, in each time step an iteration has to be performed employing, e.g., a Newton iteration scheme as given in Fig. 2.

2.2   Boundary Element Method

In the small displacement theory for homogeneous, isotropic, and linear elastic materials, the equation of motion (1) can be rewritten as

$$(c_1^2 - c_2^2)u_{i,ij} + c_2^2 u_{j,ii} = \frac{1}{\rho} b_j \tag{7}$$

where a Cartesian coordinate system ($i = 1, 2, 3$) is used. c_1 and c_2 are the velocities of the pressure and shear waves, respectively. They are defined as

$$c_1^2 = \frac{E(1-\nu)}{\rho(1+\nu)(1-2\nu)}, \quad c_2^2 = \frac{E}{2\rho(1+\nu)} \tag{8}$$

where E indicates Young's modulus, ν the Poisson's ratio, and ρ the material density.
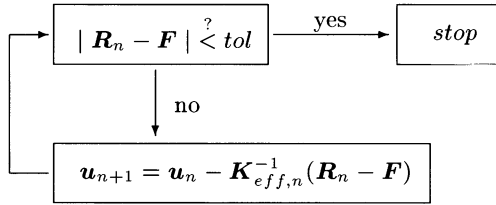
n: number of the current step in the Newton-Raphson iteration
R_n: vector containing the sum of the nodal forces corresponding to the
            element stresses at step n and the dynamic part of the equiblirium
F: global load vector
u_n: vector of the nodal displacements at step n
$K_{eff,n}$: effective stiffness matrix at step n

Fig. 2. Solution scheme for uncoupled nonlinear FEM calculations

Assuming zero body forces b_j and zero initial conditions, the equation (7) can be reduced to an integral equation of the form

$$d_i u_i(\xi, t) = \int_T \left[\int_\Gamma (u_j^* s_j - s_j^* u_j) d\Gamma \right] d\tau \tag{9}$$

in which $u_j(x, t)$ and $s_j(x, t)$ represent the displacements and tractions at the body surface Γ at time t. $u_i(x, t)$ is the displacement at an arbitrary point ξ located either on the smooth surface ($d_i = 0.5$) or in the interior of the body ($d_i = 1.0$). The fundamental solutions according to the displacements and tractions are marked by a star (*). More details can be found, e.g., in [10] and [28] .

For the numerical solution of equation (9), a discretization of the body surface with boundary elements as well as a discretization of the observation time by equal time increments Δt is needed. Collocation at each boundary node and at all time steps finally leads to a system of equations [10]

$$U^{(1)} s^{(m)} = T^{(1)} u^{(m)} + \sum_{k=1}^{m-1} \left[T^{(m-k+1)} u^{(k)} - U^{(m-k+1)} s^{(k)} \right] \tag{10}$$

which is valid for each time step m. $U^{(i)}$ and $T^{(i)}$ are influence matrices, which contain integral terms, evaluated over each boundary element and over the time step i. The vectors $u^{(k)}$ and $s^{(k)}$ contain the according displacements and tractions, respectively.

To obtain consistency between the FE and BE formulations, the boundary tractions s have to be transformed by means of a matrix M such that

$$F = Ms \tag{11}$$

where the vector $\boldsymbol{F}$ contains resultant nodal forces. For each boundary element Γ_e the elements m_{ij} of the transformation matrix $\boldsymbol{M}$ can be calculated using the shape functions N_i (see, e.g., [10] and [3])

$$\boldsymbol{M}_{ij} = \int_{\Gamma_e} N_i N_j d\Gamma. \tag{12}$$

Using equation (11), a global nodal force vector can be formulated which relates the BE nodal forces to the displacements at the time step m and a sum of the influence of all previous time steps ($k = 1, 2, ..., m - 1$):

$$\boldsymbol{F}^{(m)} = \boldsymbol{M} \left[\boldsymbol{U}^{(1)}\right]^{-1} \boldsymbol{T}^{(1)} \boldsymbol{u}^{(m)} +$$

$$\boldsymbol{M} \left[\boldsymbol{U}^{(1)}\right]^{-1} \sum_{k=1}^{m-1} \left[\boldsymbol{T}^{(m-k+1)} \boldsymbol{u}^{(k)} - \boldsymbol{U}^{(m-k+1)} \boldsymbol{M}^{-1} \boldsymbol{F}^k\right]$$

$$= \boldsymbol{K}_B \boldsymbol{u}^{(m)} + \sum_{k=1}^{m-1} \boldsymbol{H}_k^{(m)}. \tag{13}$$

2.3   Coupling Strategy

The coupling of those parts of the considered body, which are discretized by finite elements, and other parts, modeled by boundary elements, can be done by satisfying the continuity conditions along the FE/BE interfaces, namely the displacements of the FE and the BE nodes have to be equal along common interfaces and, in addition, equilibrium has to be ensured.

At the interfaces, the nodal forces resulting from the BE equation (13) can be treated as additional loads in each iteration. Consequently, the coupling algorithm can be obtained by modifying the FE algorithm given in Fig. 2. This yields the new procedure as depicted in Fig. 3. Since nonlinearities may occur only in the FE-subdomains, the geometrical linearity is checked along the interfaces by observing the strains at the interface nodes.

Assembling the FE- and BE submatrices, one obtains a coupled system of equations (14),

$$\underbrace{\begin{bmatrix} \boldsymbol{K}_{eff,n} & \\ & \boldsymbol{K}_B \end{bmatrix}}_{\boldsymbol{K}_{eff}^{(m)}} \underbrace{\begin{bmatrix} \boldsymbol{u}_F^{(m)} \\ \boldsymbol{u}_{BF}^{(m)} \\ \boldsymbol{u}_B^{(m)} \end{bmatrix}}_{\boldsymbol{u}^{(m)}} = \underbrace{\begin{bmatrix} \boldsymbol{F}_F^{(m)} \\ \hline \boldsymbol{0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{0} \\ \hline \boldsymbol{F}_B^{(m)} - \sum_{k=1}^{m-1} \boldsymbol{H}_k^{(m)} \end{bmatrix}}_{\boldsymbol{R}^{(m)}} \tag{14}$$

where the global system matrix $\boldsymbol{K}_{eff}^{(m)}$ is, due to the BE part, not symmetric and not positive definite anymore. A special solution strategy has been developed to handle these systems in an efficient way.

$$\boxed{\mid \boldsymbol{R}_n - \boldsymbol{F} + \boldsymbol{K}_B \boldsymbol{u}_n + \sum_{k=1}^{m-1} \boldsymbol{H}_k^{(m)} \mid \overset{?}{<} tol} \xrightarrow{\text{yes}} \boxed{stop}$$

no

$$\boxed{\boldsymbol{u}_{n+1} = \boldsymbol{u}_n - (\boldsymbol{K}_{eff,n} + \boldsymbol{K}_B)^{-1}(\boldsymbol{R}_n - \boldsymbol{F} + \boldsymbol{K}_B \boldsymbol{u}_n + \sum_{k=1}^{m-1} \boldsymbol{H}_k^{(m)})}$$
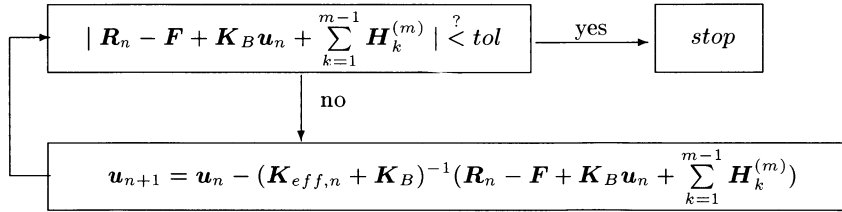
m: number of the current step
$\boldsymbol{H}$:   vector containing the influence of all previous time steps
$\boldsymbol{K}_B$:   effective stiffness matrix (BEM)

Fig. 3. Solution scheme for the coupled nonlinear FEM/BEM calculations

2.4   Computational Aspects

The new coupling procedure discribed in the privious sections can be devided in the following steps:

1. computation of the BEM influence matrices
2. transformation of the influence matrices with $\boldsymbol{M}$ (equations 12 and 13)
3. assembling of the coupled stiffness matrix $\boldsymbol{K}_{eff}^{(m)}$
4. solving of the system of equations in order to obtain the nodal displacements
5. computation of the secondary unknowns, e.g., the stresses.

In most cases, the computations of step 1 and step 4 are the most expensive one compared to the the others.
For the computation of the influence matrices (step 1) and their transformation (step 2), a considerable number of operations exists, which needs to be repeated using various parameters. An essential speed-up could be obtained by implementing parallel processing of the algorithms of the steps 1 and 2 using MPI (Message Passing Interface) [30]. In the case of four examples, whose dimensions are given in Table 1, the improvement is schown in Fig. 4. Choosing a suitable number of processors, a CPU usage of the processors close to 100 percent can be reached.
The resulting stiffness matrix of the coupled system is nonsymmetric and not positive definit, however, sparsely populated. Therefore special solution algorithms are used for the inversion of the system of equations (step 4). A significant improvement with respect to the computation time could be obtained by employing the solver provided in the SuperLu subroutine library [41].
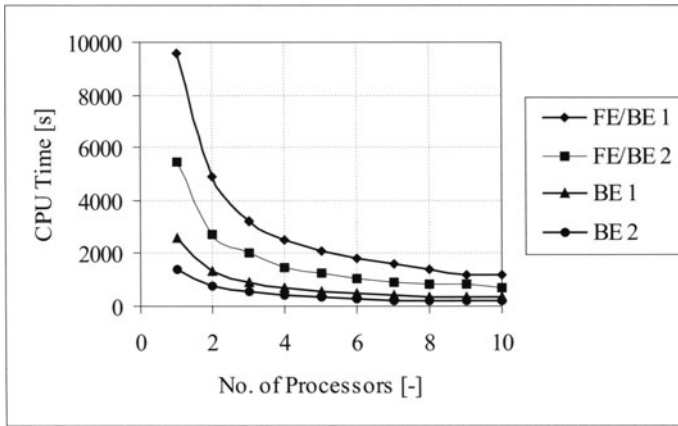
Fig. 4. CPU-time for the computation of influence matrices

Table 1. Dimensions of the four examples (coupled/uncoupled)

System	Dim. of influence matrices	No. of influence matrices
FE/BE 1	1107	51
FE/BE 2	1107	28
BE 1	675	51
BE 2	675	28

3   Numerical Examples

According to the development steps of the new model, the numerical examples documented next are distinguished in two- and three-dimensional investigations.

3.1   Two-Dimensional Systems

The two-dimensional model provides a very helpful approach to gain a first idea of the dynamical behavior of a complicated track system. It can be used under the assumption that the loading condition and the system do not change in one direction (perpendicular to the investigated plane). Two representative examples are discussed, namely a railway track on a halfspace and a track in a tunnel. The latter is thought to demonstrate the flexibility of the new approach.

Railway Track on a Halfspace. The example deals with the nonlinear behavior of a railway track on a halfspace. The system consists of three sub-

regions: the sleepers (region I, discretized with finite elements), the ballast including the subgrade (region II, discretized with finite elements), and the undisturbed soil (region III, discretized with finite elements and with boundary elements). The complete geometry and its discretization can be found in Fig. 5.
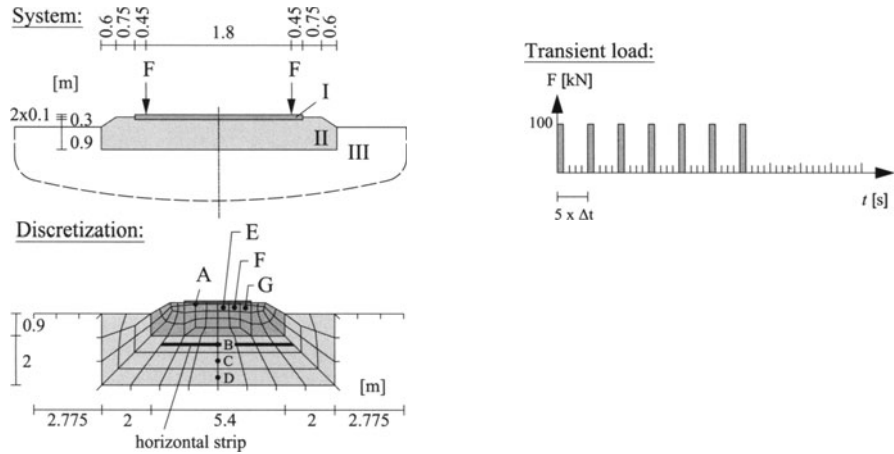
Materialparameters:

Sleeper:

Bulk modulus: $\kappa = 6.68 \times 10^6 \ kN/m^2$
Lame constant: $\mu = 2.63 \times 10^6 \ kN/m^2$
density: $\rho_0 = 2.2 \ t/m^3$

Ballast:

Young's modulus: $E = 1.28 \times 10^5 \ kN/m^2$
Poisson's ratio: $\nu = 0.25$
$\rho_0 = 1.7 \ t/m^3$
initial yield stress: $Y_0 = 10 \ kN/m^2$

Soil
$E = 5.7 \times 10^7 \ kN/m^2$
$\nu = 0.25$
$\rho_0 = 1.7 \ t/m^3$
Damage in near field: $\alpha = 1 \ and \ \beta = 0.5$

Fig. 5. Railway track on a halfspace

In particular, also the parameters needed afterwards for the different material laws are depicted in Fig. 6. Throughout the study, the sleepers (region I) are represented by a Neo-Hooke material, while in the case of the ballast and the subgrade (region II) an elastoplastic material with isotropic and kinematic hardening is used as proposed by von Mises (see, e.g. [5]). Using special finite elements for parts of the undisturbed soil (region III, FE model), also the investigation of local damages is possible. In this example, the damage model given by Simo [35] is used. It is based on an elastic model which is extended by a damage function depending on two additional material parameters α and β. The remaining part of the soil (region III, BE model) is assumed to be linear elastic.
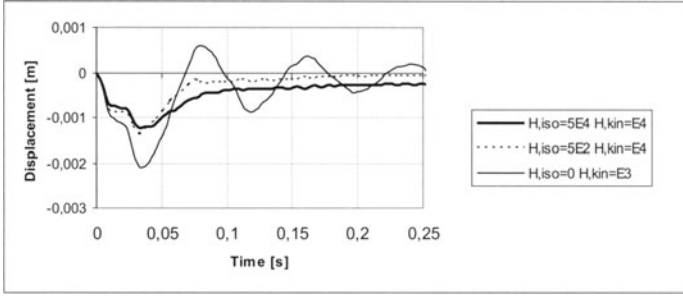
Fig. 6. Time history of the vertical displacements at point A

The system is subjected to two vertical loads acting in time as a sequence of impulses of $F = 100\ kN$, which each has a duration of one Δt (Fig. 5). The size of the time step Δt has been chosen such that the pressure wave in the undisturbed soil is able to cross one element of the BE discretization within one step.

Fig. 6 shows the time history of the vertical displacements at point A located below the sleeper (see Fig. 5). Different combinations of the hardening parameters H_{iso} (isotropic) and H_{kin} (kinematic) are considered. The maximum displacement occurs for $H_{iso} = 0$. For this case also a positive displacement can be observed, which means that some tension occurs in the system (i.e., the model is not valid anymore). Generally the displacements reduce to a constant value as soon as the load is set to zero.
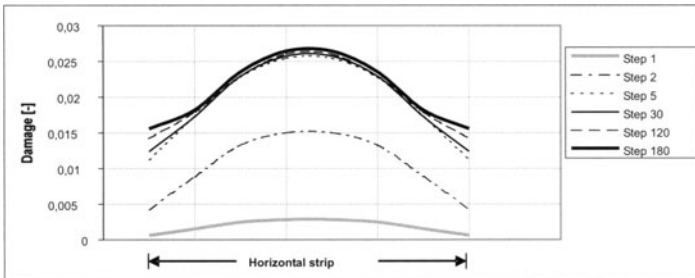
Fig. 7. Distribution of the damage in a horizontal strip directly underneath the ballast at different time steps

To investigate the distribution of the damage inside the soil material, a horizontal strip directly underneath the ballast is considered (see Fig. 5). The material parameters of the ballast are assumed to be $Y_0 = 10\ kN/m^2$ (initial yield stress), $H_{iso} = 5.10^5\ kN/m^2$ and $H_{iso} = 104\ kN/m^2$. In Fig. 7 the distribution of the damage after 180 load impulses is given. As expected, the maximum occurs at the center of the considered strip and it may be observed
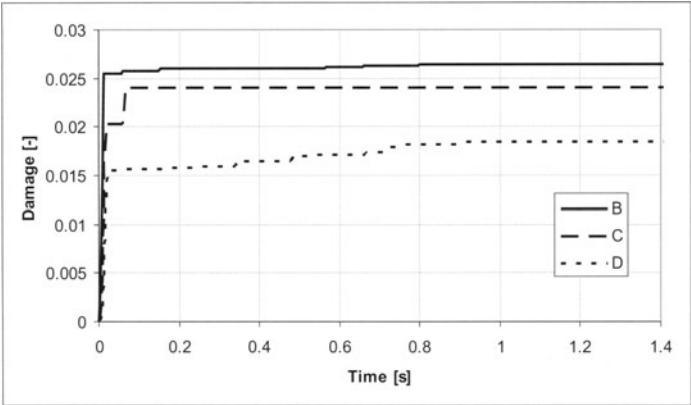
Fig. 8. Time history of the damage at points B, C and D
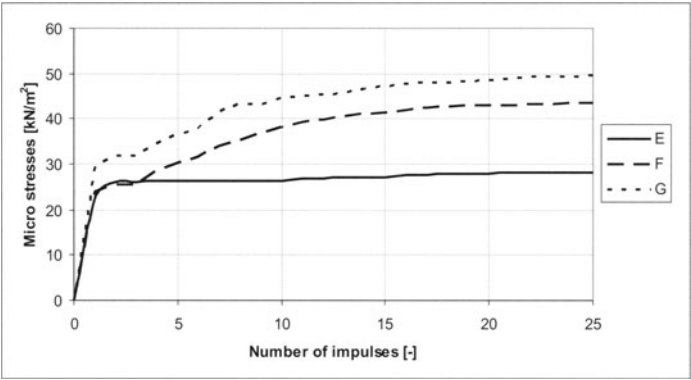
Fig. 9. Micro stresses (isotropic hardening) at points E, F and G for a different number of impulses

that the damage of the soil is converging to a final maximum value. Looking at the time histories of the damage at given points B, C and D, which are depicted in Fig. 8, the same convergence effect can be observed.

Finally, the yielding behavior of the ballast is investigated. As before, the micro stress, qH_{iso} resulting from the isotropic hardening, is considered. Fig 9 shows, at given points E, F and G (see Fig. 5), the dependence of this parameter from the number of load impulses. It may be seen that also in this case the hardening is approaching a final maximum value. Once this value is reached, additional load impulses do not lead to further changes.

Track in a Tunnel. In this example a tunnel which is excited by a passing train is considered. In particular, it is investigated whether the stress in the surrounding soil reaches the yielding point.

The system is given in Fig. 10. It can be divided into four subregions: the sleeper, the ballast, the tunnel liner, and the surrounding soil. The sleeper is modeled with a nonlinear elastic material law (Neo-Hooke). The ballast is assumed to have viscous properties in deviatoric direction only.

System:
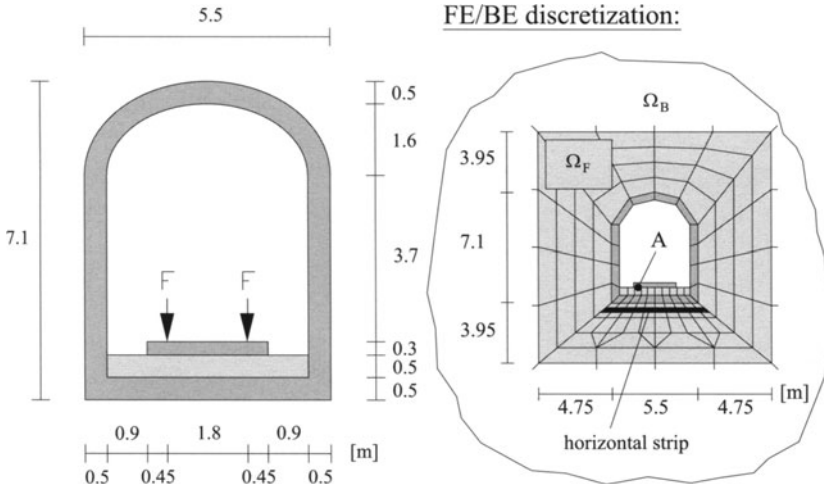
Transient load:
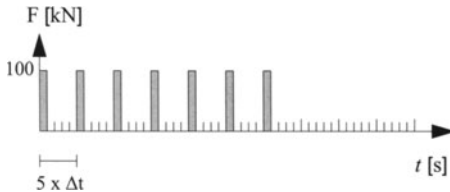
Materialparameters:

Sleeper:
Bulk modulus: $\kappa = 6.68 \times 10^6 \ kN/m^2$
Lame constant: $\mu = 2.63 \times 10^6 \ kN/m^2$
density: $\rho_0 = 2.2 \ t/m^3$

Ballast:
$\kappa = 9.11 \times 10^4 \ kN/m^2$
$\mu = 5.47 \times 10^4 \ kN/m^2$
$\rho_0 = 1.7 \ t/m^3$
$\tau = 10^{-2} \ s$
$\gamma = 0.5$

Tunnel liner:
Young's modulus: $E = 3 \times 10^7 \ kN/m^2$
Poisson's ratio: $\nu = 0.16$
$\rho_0 = 2.5 \ t/m^3$

Soil:
$E = 5.7 \times 10^7 \ kN/m^2$
$\nu = 0.25$
$\rho_0 = 1.7 \ t/m^3$
In near field:
initial yield stress: $Y_0 = 10 \ kN/m^2$

Fig. 10. Discretization of the tunnel system located in a fullspace

The material damping can be described by two additional parameters, namely the relaxation time τ and a parameter γ, which occur in the evolution equa-

tion of the internal variable of the viscosity. Details are given, e.g., [35]. The liner consists of reinforced concrete for which a St. Venant material law is assumed.

In the vicinity of the tunnel, the surrounding soil is modeled by finite elements. The remaining fullspace is represented by boundary elements. In order to be able to investigate the hardening process in the soil around the tunnel, an elastoplastic material law as suggested by von Mises (see [5]) is used in the finite element subregion. It also contains isotropic hardening. In the boundary element subregion, where mainly the wave propagation to infinity shall be taken into account, an elastic soil material is assumed.

To study the development of hardening effects in the ballast and the soil, the vertical load representing the passing train is approximated by 15 rectangular impulses (see Fig. 10). Each impulse has a duration of two time steps Δt.

First, the vertical displacements at point A, located below the sleeper, is investigated (see Fig 11). Different values of the parameter H_{iso}, which describes the isotropic hardening, are considered. It should be noted that with increasing H_{iso} the yielding effects are getting more significant which results in a decrease of the plastic deformations. Also a certain time shift between the load impulses and the occurrence of the maximum (delay) may be observed.
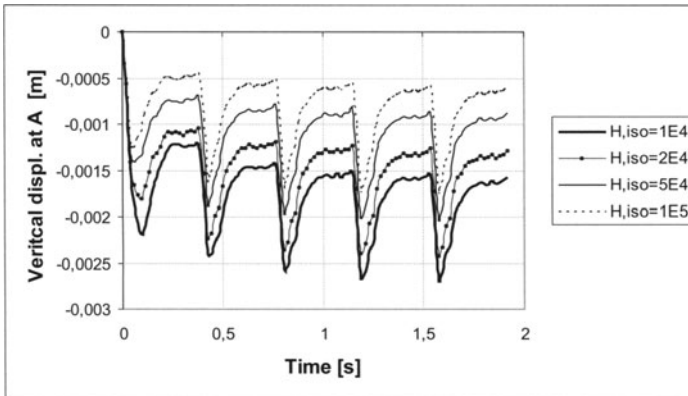
Fig. 11. Influence of the isotropic hardening parameter H_{iso} of the soil on the vertical displacement at A

Second, the distribution of the micro stresses qH_{iso} (hardening) in a horizontal strip directly underneath the tunnel (see Fig. 10) is determined, assuming a constant $H_{iso} = 105 \ kN/m^2$. Note that the micro stresses are a common quantity to describe the state of the material hardening; q is a hardening variable. In the present case, only one load impulse of time duration $2\Delta t$

has been prescribed. The result is given in Fig. 12. It may be observed that the maximum of the hardening occurs below the center of the tunnel where the influence of the two vertical loads is superimposed. Close to the edges of the tunnel liner, the hardening is increasing as well. This is due to the stress concentration at these points.
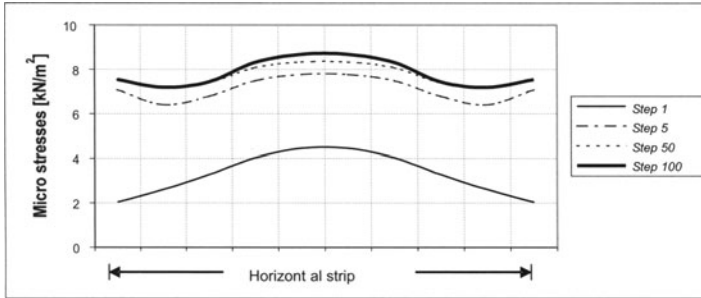
Fig. 12. Distribution of the micro stresses (hardening) in a horizontal strip directly underneath the tunnel liner at different time steps

3.2   Three-Dimensional Systems

In many cases, e.g., if a moving train shall be considered, a three-dimensional element model of the track system is needed. Of course, due to limitations arising from the available computer power, only a small part of the complete track can be discretized. Usually, a limited number of sleepers is considered, mainly in order to take into account local three-dimensional effects. Again, two representative examples are investigated: a rigid sleeper on a halfspace and on a more sophisticated soil model. It should be mentioned that an additional example, namely a system consisting of two sleepers, can be found in the contribution of Schmid and Friedrich in this book.

Rigid Sleeper on an Elastic Halfspace (Verification). Consider the three-dimensional system given in Fig. 13. It consists of a rigid sleeper of $20 \times 26 \times 260$ [cm], $\rho = 2145 \ kg/m^3$, which is placed on an elastic halfspace, whose material properties are $E = 1.7 \times 10^5 \ kN/m^2$, $\nu = 0,25$, and $\rho = 1700 \ kg/m^3$. While the sleeper is modeled with $2 \times 2 \times 20$ quadratic finite elements, the halfspace is represented by 28×10 quadratic boundary elements. The system is loaded by two vertical forces located at the position of the rails. They act in time as a rectangular impulse over a time duration of 50 time steps $\Delta t = 0.00065s$.

It should be mentioned that the configuration chosen here has been defined within the context of a benchmark of different computational models developed during the course of the priority program. More details about the
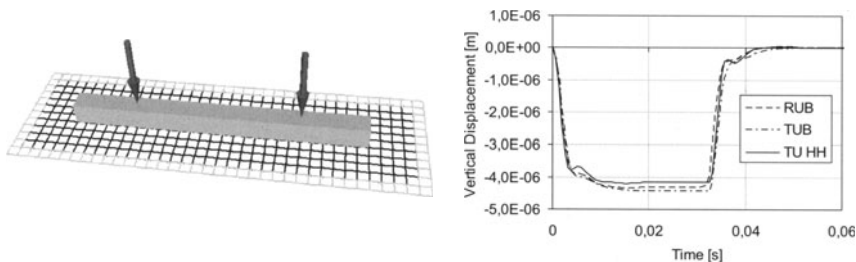
Fig. 13. Discretization of the rigid sleeper on a halfspace (left) and a comparison of different computational models (right)

benchmark can be found in the contribution by Rücker, et al. in this book. In order to verify the new model developed above, its results (TU HH) are compared to solutions calculated by Schmid et al. (RUB) and Savidis et al. (TUB). Whereas Schmid and his group are using a BEM approach in the frequency domain (here transformed to the time domain), Savidis and his co-workers are coupling finite elements with the thin-layer approach [18]. Fig. 13 shows the transient vertical displacement of the rigid sleeper. In the loading phase, the system reaches its static state after about $0.012\ s$ and returns to the initally undeformed situation again after unloading. It can be observed that the results obtained with the three different models agree very well.

It should be pointed out that the different methodologies also have been compared to measurements. The respective results are discussed in this book by Lammering and Plenge and shall not be repeated here.

Rigid Sleeper on a nonlinear soil model. In order to demonstrate the applicability of the new approach in the case of a more realistic soil model, a rigid sleeper of $20 \times 25 \times 250$ [cm], $\rho = 2145\ kg/m^3$, is placed on a halfspace as before. However, the soil in the vicinity of the sleeper, in a subdomain of $75 \times 75 \times 300$ [cm], is modeled using an elastic-plastic material law with a yield surface suggested by Drucker-Prager [5]. The effects of a kinematic hardening are also included in this model. The soil outside the subdomain underneath the sleeper, i.e. the halfspace, is assumed to behave linearly elastic. Its material parameters are the same as in the benchmark example ($E = 1.7 \times 10^5\ kN/m^2$, $\nu = 0,25$). Within the subdomain, the elastic-plastic model necessitates additional material parameters which are the cohesion, $c = 10.5\ kN/m^2$, the angel of internal friction, $\phi = 45$, and a hardening parameter, $H = 100\ kN/m^2$.

FEM   BEM   Coupling of FEM and BEM
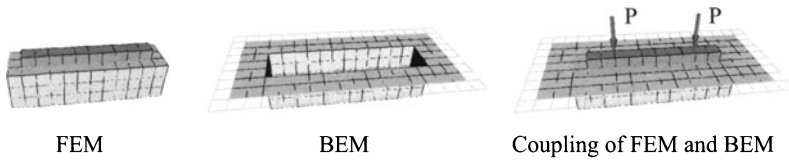
Fig. 14. Discretization of the sleeper on the nonlinear soil model

The discretization of the FEM and BEM parts are depicted in Fig. 14. The sleeper is discretized with 10 quadratic finite elements, the soil material in the near field is modeled with $3 \times 3 \times 12$ qadratic finite elements, and the remaining soil is represented by 252 quadratic boundary elements. The system is loaded, as before by two vertical forces located at the position of the rails. They act in time as four rectangular impulse of time duration of $2 \times \Delta t$ each with a period of $10 \times \Delta t$, where $\Delta t = 7.216 \times 10^{-4}$ s. The amplitude of the load is $P = 50\ kN$.

Fig. 15 (left) shows the vertical displacement of the sleeper due to the four impulses. The remaining displacement after unloading can be seen in Fig. 15 (right). The growth of the plastic strain is given in Fig. 16, where A is a point directly underneath the sleeper below the acting force. The points B and C are located on the common interface between the FE- and the BE-mesh: C vertically below A and B below the center of the sleeper. In all points, the influence of the four impulses can be observed.
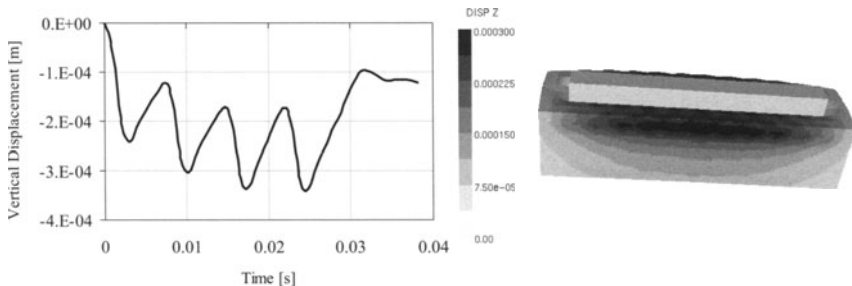
Fig. 15. Transient vertical displacement of the sleeper (left) and the remaining soil deformations after unloading (right)

4   Conclusion

Extending a linear two-dimensional FEM/BEM coupling scheme to a two- and three-dimensional formulation where nonlinearities can be taken into ac-
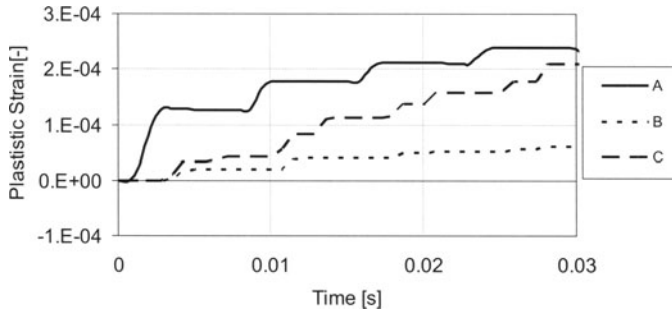
Fig. 16. Transient development of the plastic strain of three representative points A, B and C

count, the number of possible investigations with respect to railway track systems could be increased considerably.

From the numerical examples documented in this contribution, it became obvious that the new approach seems to work very well. In particular, restricting possible nonlinearities to the FE-subregions, it was shown that the discretization of rather realistic track systems can be done in a very flexible way.

In the future, emphasis will be placed on the implementation of more sophisticated soil models, for instance, as suggested by Gudehus et al.[15].

Acknowledgement

References

1. Adam M., Pflanz G., Schmid G. (2000) Two- and Three-Dimensional Transient Responses of Half-Space under Dynamic Strip Loads and Train Track Loads. Proc. of the 12th World Conference on Earthquake Engineering, Auckland, New Zealand
2. Bathe K.J. (1996) Finite element procedures. Prentice Hall. Upper Saddle River
3. Becker A.A. (1992) The boundary element method in engineering. McGraw-Hill, Berkshire

4. Beskos DE. (1997) Boundary element methods in dynamic analysis. Part II (1986-1996). Appl. Mech. Rev. **50**, 149-197
5. Chen WF. (1994) Constitutive Equations for Engineering Materials, Volumen2: Plasticity and Modeling. Elsevier, Amsterdam
6. von Estorff O., Firuziaan M. (1999) Nonlinear Dynamic Response by Coupling BEM and FEM. Proc. of the European Conference on Computational Mechanics (ECCM 99) Munich
7. von Estorff O., Firuziaan M. (2000) Coupled BEM/FEM Approach for Nonlinear Soil/Structure Interaction. Engineering Analysis with Boundary Elements, Elsevier Science
8. von Estorff O., Firuziaan M. (2000) Transient Nonlinear Behaviour of Railway Beds by FEM/BEM Coupling. EUROMECH Colloquium 409, "Dynamics and Long-Term Behaviour of Railway Vehicles, Track and Subgrade", Hannover.
9. von Estorff O., Firuziaan M. (2000) FEM and BEM for Nonlinear Soil/Structure Interaction Analyses. International Workshop WAVE 2000, Bochum
10. von Estorff O., Prabucki MJ. (1990) Dynamic response in the time domain by coupled boundary and finite elements. Computational Mechanics. **6**, 35-46
11. Firuziaan M., von Estorff O. (2002) Transient 3D Soil/Structure Interaction Analyses Including Nonlinear Effects. eurodyn 2002 (submitted)
12. Fukui T. (1987) Time marching BE-FE method in 2-D elastodynamic problems. Int. Conf. on BEM IX, Stuttgart
13. Gaul L., Schanz M. (1998) Material Damping Formulations in Boundary Element Methods. Proc. of the 16th Int. Modal Analysis Conf., Santa Barbara
14. Gaul L. (1999) Influence of Damping on Waves and Vibrations. Mechanical Systems & Signal Processing **13**, 1-30
15. Gudehus G., Kolymbas D. (1979) A constitutive law of the rate type for soil. 3rd Int. Conf. Num. Meth. Geomech. (Ed. W. Wittke), Aachen, Balkema, 319-329
16. Kausel E., Roësset J.M. (1975) Dynamic stiffness of circular foundations, Journal of Engineering Mechanics Division, ASCE **101**, 771-785
17. Kausel E. (1988) Local transmitting boundary, Journal of Engineering Mechanics, ASCE , 1011-1027
18. Kausel E. (1994) Thin-Layer-Method: Formulation in the Time Domain, Int. J. for Num. Methods in Eng., Vol. 37, 927-941.
19. Kayupov M.A., Bulgakov V.E., Kuhn G. (1998) Efficient Solution of 3-D Geomechanical Problems by Indirect BEM Using Iterative Methods. Int. J. for Numerical and Analytical Methods in Geomechanics **22**, 983-100
20. Knothe K., Wu Y. (1998) Receptance Behavior of Railway Track and Subgrade. Archive of Applied Mechanics **68**, 457-470
21. Lefeuve-Mesgouez G. (1999) Ground Vibration due to a high speed moving harmonic load. Proc. of the 4th European Conference on Structural Dynamics, EURODYN '99, Prag
22. Lima L.T., (1999) A Two-dimensional Conjugated Infinite Element Method for Elastodynamics: Formulation and Validation in the Frequency Domain. Ph.D. Thesis, UCL Universit catholique de Louvain, Louvain-la-Neuve
23. Luo Y., Yin H., Hua C.(1996) Dynamic Response of Railway Ballast to the Action of Trains Moving at Different Speeds. Proc. of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit **210**, 95-101
24. Lysmer J. (1970) Lumped mass method for rayleigh waves. Bulletin of the Seismological Society of America. **60**, 89-104

25. Lysmer J., Kuhlemeyer R.L. (1969) Finite dynamic model for infinite media. Journal of Engineering Mechanics Division, ASCE **95**, 859-877

26. Mansur W.J. (1983) A time-stepping technique to solve wave propagation problems using boundary element method. Ph.D. Thesis. University Southampton

27. Peplon A.T., Jones C.J.C., Petyt M. (1999) Surface Vibration Propagation over a Layered Elastic Halfspace with Inclusions. Applied Acoustics **56**, 283-296

28. Pflanz G. (2001) Numerische Untersuchung der elastischen Wellenausbreitung infolge bewegter Lasten mittels der Randelementmethode im Zeitbereich. Fortschritt-Berichte VDI Reihe 18, Nr. 265, VDI-Verlag, Düsseldorf

29. Pflanz G., Hubert W., Schmid G. (1999) Erschütterungsausbreitung durch Schienenverkehr. Tagungsband des 11. Forums Bauinformatik, TU Darmstadt

30. Rauber T., Rüger G. (2000) Parallele und verteilte Programmierung. Springer Verlag, Berlin

31. Ripke B.(1995) Hochfrequente Gleismodellierung und Simulation der Fahrzeug-Gleis-Dynamik unter Verwendung einer nichtlinearen Kontaktmechanik. Fortschritt-Berichte VDI Reihe 12, Nr. 249, VDI-Verlag, Düsseldorf

32. Ripke R. (1999) Dynamik und Langzeitverhalten des Schotters - Anforderungen an die Simulationsmodelle. In: Schotteroberbau - Mechanische Modellierung, Laborversuche und die Praxis. Symposium des Interdisziplinren Forschungsverbundes Bahntechnik, Berlin

33. Savidis S.A., Bode C., Hirschauer R. (2000) Three-Dimensional Structure-Soil-Structure Interaction Under Seismic Excitation with Partial Uplift. Proc. of the 12th World Conference on Earthquake Engineering 12WCEE, New Zealand

34. Savidis S.A., Hirschauer R. (1998) Dynamic soil-structure interaction of adjacent structures. Proc. of the Int. Conf. on Soil-Structure Interaction in Urban Civil Engineering, Darmstadt

35. Simo J.C. (1987) On a Fully Three-Dimensional Finite-Strain Viscoelastic Damage Model: Formulation and Computational Aspects. Computer Methods in Applied Mechanics and Engineering **60**, 153-173

36. Turek J. (1995) Non-linear Response of the Track. Vehicle System Dynamics **24**, 265-279

37. Waas G. (1972) Linear two-dimensional analysis of soil dynamics problems in semi-infinite layered media. Ph.D. Thesis. University of California

38. Wolf J.P., Song C.h. (1999) The Guts of Dynamic Soil-Structure Interaction. Proc. of the Int. Symp. On Earthquake Engineering ISEE99, Budva, Montenegro

39. Wolf J.P. (1994) Foundation Vibration Analysis Using Simple Physical Models. Prentice-Hall, Englewood Cliffs

40. Wolf J.P., Song C.h. (1997) Finite-Element Modelling of Unbounded Media. John Wiley and Sons, Chichester

41. Xiaoye S.L., Demmel J.W. (1998) Making Sparse Gaussian Elemination Scaleable by Static Pivoting. Proceeding of SC 98 Conference, Orlando, Florida

Dynamic Behavior of Railway Track Systems Analyzed in Frequency Domain

Klaus Friedrich and Günther Schmid

Ruhr-University Bochum, Fakultät für Bauingenieurwesen,
Universitätsstr. 150, 44780 Bochum, Germany

Abstract. An analysis procedure is developed in order to describe the system-dynamics of railroad track, substructure and soil. It allows to model the dynamic behavior of rails, elastic pads, sleepers, railroad earthworks and subsoil. By means of the developed program the dynamic features of various track-models under harmonic load as well as under moving load can be analyzed. A model is developed for a passing train which allows the simulation of the influence of the moving load by means of a frequency- and time-domain transformation. Rails, sleepers and elastic pads are modelled by finite elements while the system-components extending to infinity, such as rigid track, ballast and subsoil, for which the wave radiation is essential, are described with boundary elements. The developed model is validated by a comparison with other numerical models in the frame of a benchmark test.

1 Introduction

Today's increasing ground vibration and stress levels on railroad track, substructure and subsoil due to the velocity of modern high speed trains call for the development of optimized track-systems. Features of the track which are subjected to dynamic or moving loads have to be analyzed regarding the dynamic behavior. Additionally, economically suitable systems with high endurance and low maintenance costs need to be built. The simulation of the passing train with numerical models is a suitable means in order to display the dynamic behavior next to extensive measurements, as in computational simulations different track-geometries, material parameters and the layered soil can be varied easily. In this way the influence of the most diverse parameters on the dynamic behavior of the track can be studied. It therefore allows the optimization of the railroad track and substructure without the necessity to build extensive test tracks. Suitable mathematical formulations for the various subsystems were chosen in order to display the dynamic features of the individual components in the model. Subsystems such as the infinitely extended soil and those parts of the track for which the wave radiation is essential, are modelled by boundary elements while parts of the superstructure such as rails, sleepers and elastic pads are modelled by finite elements. At the nodal points of the common surface the subsystems are coupled by means of the substructure technique. The large number of recent publications in this discipline emphasizes its importance. A numerical analysis of the wave propagation caused by loads moving with sub- and supercritical velocity using a

time-domain approach is presented in [15]. Other investigations focussing on railway dynamics and moving loads can be found in [14], [11] and [12]. In order to meet the requirements of high speed trains, an elastically supported track system with ladder sleepers is developed in Japan [18]. Discrete and continuous models for the railway track dynamics subjected to high frequency vertical excitation can be found in [8]. Extended analytical and numerical examinations on the dynamic interaction of rails, pads, sleepers and subsoil are given in [10] and [19].

The objective of this project is the integration of the different numerical tools in one program which satisfies the aspects mentioned above. It facilitates to carry out the analysis under harmonic load as well as under moving load. Although the numerical method is also suitable for horizontal dynamics, the investigations done in this project are focussing on the more important vertical dynamics of track-subsoil interaction. The results are evaluated and compared in a benchmark test containing a comparison with measured data.

2   Analysis Procedure

In this chapter the basic aspects of the numerical analysis procedures which are implemented in the developed program are presented.

2.1   Boundary Element Method (BEM)

The boundary element method is based on the fundamental solution of the differential equation of motion for a homogeneous, isotropic, linear-elastic continuum [6]. The equation, also known as Lamé-Navier equation, can be obtained by the following relations: The equilibrium leads to

$$\sigma_{ij,j} + f_i = \varrho \ddot{u}_i \,, \tag{1}$$

where σ_{ij} , f_i and u_i are the components of stress, volume force and displacement, respectively and $\varrho \, \ddot{u}_i$ is the force due to inertia. The commas and dots indicate space and time derivatives. The material law for elastic material is given as

$$\sigma_{ij} = \lambda \varepsilon_{kk} \delta_{ij} + 2\mu \varepsilon_{ij} \,, \tag{2}$$

where λ and μ are the Lamé constants and δ_{ij} is the Kronecker delta. The kinematic compatibility is expressed as

$$\varepsilon_{ij} = \frac{1}{2} (u_{i,j} + u_{j,i}) \,. \tag{3}$$

Introducing (2) and (3) into (1) leads to the wave equation:

$$\varrho \left[\left(c_p^2 - c_s^2 \right) u_{i,ik} + c_s^2 u_{k,ii} - \ddot{u}_k \right] = -f_k \,, \tag{4}$$

where $c_p = \sqrt{\frac{\lambda+2\mu}{\varrho}}$ and $c_s = \sqrt{\frac{\mu}{\varrho}}$ are the pressure and shear wave velocities, respectively. In order to obtain the wave equation in frequency domain the application of the Fourier Transform leads to

$$\varrho[(c_p^2 - c_s^2)\tilde{u}_{i,ik} + c_s^2\tilde{u}_{k,ii} + \omega^2\tilde{u}_k] = -\tilde{f}_k . \tag{5}$$

In (5) $\tilde{u}_k$ and $\tilde{f}_k$ can be interpreted as the complex displacement and force amplitudes, respectively of a harmonic motion with circular frequency ω.
When body forces are assumed to be zero and homogeneous initial conditions apply, the boundary integral equation can be obtained with the help of Betty's law as follows:

$$c_{ik}u_i^\star = \int_\Gamma u_{ik}^\star t_i d\Gamma - \int_\Gamma t_{ik}^\star u_i d\Gamma , \tag{6}$$

where $\Gamma = \Gamma_1 \cup \Gamma_2$ is the boundary of the domain Ω, $t_{ik}^\star(x,\xi,\omega)$ and $u_{ik}^\star(x,\xi,\omega)$ are the fundamental solutions for the traction and displacement components in direction i at point x due to a unit load in direction k at point ξ. For smooth boundaries the elements of the boundary matrix c_{ik} are equal to δ_{ik} if $\xi \in \Omega$ and $0.5\,\delta_{ik}$ if $\xi \in \Gamma$. On the boundaries Γ_1 and Γ_2 the displacement boundary conditions $\bar{u}_i(x,\omega)$ and the traction boundary condition $\bar{t}_i(x,\omega)$ are prescribed, respectively. Using a complex young's modulus $E = E(1 + i\omega\eta)$ with viscous damping coefficient η, the frequency domain formulation allows to incorporate damping. The discretization of the boundary leads to the algebraic form of the integral equation. In matrix notation this can be written as:

$$U\,t = T\,u , \tag{7}$$

where u and t are the complex frequency dependent displacements and tractions of all nodal points at the boundary and U and T are the influence matrices. The time dependent displacements and tractions can be obtained by Inverse Fourier Transform as done in Sect. 3.3.

2.2 Finite Element Method (FEM)

Using the Principle of Virtuel Displacements, the equation of motion of the finite element method in the frequency domain can be obtained as:

$$-\omega^2 M u + K\,u = P , \tag{8}$$

where

$$M = \int_\Omega \varrho N^T N d\Omega ,$$

$$K = \int_\Omega B^T D B d\Omega \tag{9}$$

are the mass and stiffness matrices, respectively, and

$$P = \int_\Gamma N^{\mathrm{T}} t d\Gamma + \int_\Omega N^{\mathrm{T}} f d\Omega \qquad (10)$$

are the nodal forces. The Matrix B contains the relation between strain and displacements, D describes the connection between stress and strain and N contains the shape functions. Equation (8) can be written as

$$P = K^{\mathrm{FEM}} u , \qquad (11)$$

with

$$K^{\mathrm{FEM}} = K - \omega^2 M . \qquad (12)$$

In case of beam elements which are used in this paper, the mass and stiffness matrix are obtained by an analytical integration.

2.3   Coupling of BEM and FEM

A coupling of finite elements and boundary elements allows the examination of structures consisting of domains with different dynamic behavior and additionally it enables to benefit from the characteristics of each particular formulation. In this project those parts of the structure whose dynamic behavior essentially depends on the wave radiation like subsoil, ballast or rigid track are modelled with boundary elements while sleepers, rails and elastic pads are modelled with finite elements (see Fig.1). The sleepers and rails are discretized with Timoshenko beam elements containing shear deflection. The structure of the corresponding element mass and stiffness matrix which can be obtained from (9) is given in [17]. The elastic pads are discretized with linear viscous spring-damper elements. The degrees of freedom of each finite element node are:

$$u_i^{\mathrm{FEM}} = \begin{bmatrix} u_x \\ u_y \\ u_z \\ \varphi_x \\ \varphi_y \\ \varphi_z \end{bmatrix}_i^{\mathrm{FEM}} . \qquad (13)$$

The track and the soil are modelled with constant boundary elements with three degrees of freedom at each node:

$$u_i^{\mathrm{BEM}} = \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix}_i^{\mathrm{BEM}} . \qquad (14)$$
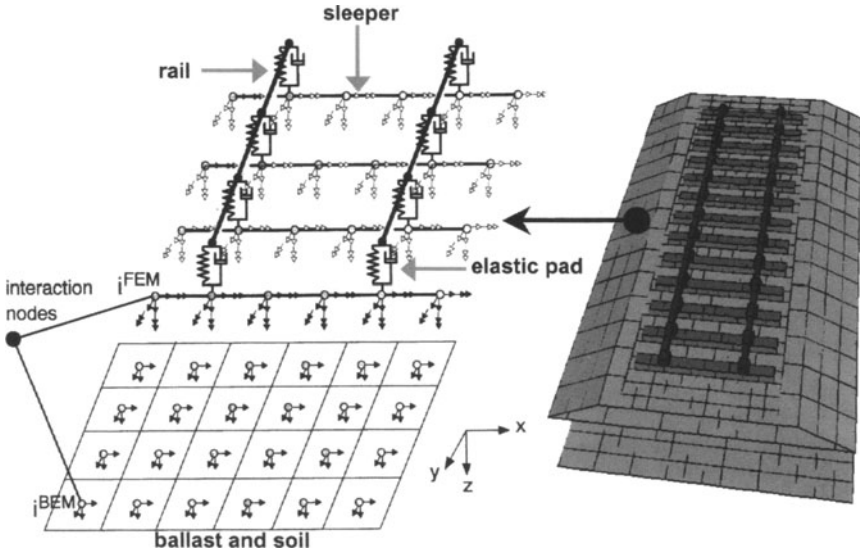
Fig. 1. Model of the track-soil system

The coupling conditions for the displacements at the contact points between the finite element region and the boundary element region are:

$$u_{x_i}^{\text{BEM}} = u_{x_i}^{\text{FEM}} \, ,$$
$$u_{y_i}^{\text{BEM}} = u_{y_i}^{\text{FEM}} \, ,$$
$$u_{z_i}^{\text{BEM}} = u_{z_i}^{\text{FEM}} \, . \tag{15}$$

In order to obtain a stiffness matrix of the boundary element region, (7) has to be transformed to relate nodal displacements with nodal forces

$$\boldsymbol{K}^{\text{BEM}} \boldsymbol{u} = \boldsymbol{P} \tag{16}$$

and

$$\boldsymbol{P} = \boldsymbol{A} \boldsymbol{t} \, . \tag{17}$$

The dynamic stiffness matrix $\boldsymbol{K}^{\text{FEM}}$ of the rails, the sleepers and the pads, and the dynamic stiffness matrix $\boldsymbol{K}^{\text{BEM}}$ of the boundary element part have to be assembled at the interaction nodes of the surface. To do this the rank of the boundary element matrix has to be expanded to the degrees of freedom of the finite element matrix inserting zeros at the corresponding rotational degrees of freedom. The equation of the coupled system can be written in expanded form as:

$$\begin{bmatrix} \boldsymbol{K}_{FF}^{FEM} & \boldsymbol{K}_{FC}^{FEM} & \boldsymbol{0} \\ \boldsymbol{K}_{CF}^{FEM} & \boldsymbol{K}_{CC}^{FEM} + \boldsymbol{K}_{CC}^{BEM} & \boldsymbol{K}_{CB}^{BEM} \\ \boldsymbol{0} & \boldsymbol{K}_{BC}^{BEM} & \boldsymbol{K}_{BB}^{BEM} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_F \\ \boldsymbol{u}_C \\ \boldsymbol{u}_B \end{bmatrix} = \begin{bmatrix} \boldsymbol{P}_F \\ \boldsymbol{P}_C \\ \boldsymbol{P}_B \end{bmatrix} , \qquad (18)$$

where the indices F and B indicate the finite element and the boundary element area and the index C indicates the coupled part of the structure as sketched in Fig. 1. General descriptions on the coupling of finite elements and boundary elements are presented in [1]. Similar coupling algorithms especially for the track-subsoil interaction can be found in [21], [3] and [12]. A numerical solution for structure-soil interaction problems based on Green's functions is described in [2].

3   Numerical Results

In order to obtain information about the dynamic behavior of the railroad track and the subsoil, ballasted tracks and rigid tracks on layered soil are analyzed. The examination is performed with tracks subjected to a fixed harmonic load as well as to a moving load to contain information about the influence of the train velocity.
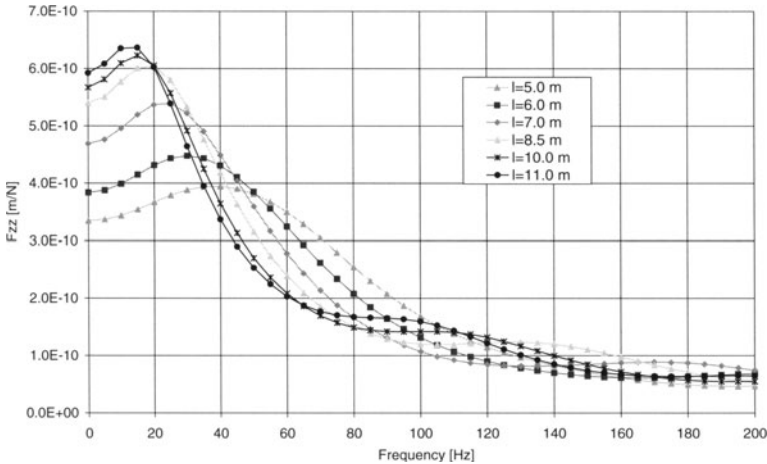
Fig. 2. Influence of the model length l

3.1   Model Size

As the area of discretization plays an important role in the boundary element method when the domain extends to infinity and a full-space fundamental

solution is used, a first investigation is done on the influence of the length of the discretized model. In this example a harmonic unit load is directly applied on the concrete layer of a rigid track on the half-space and the magnitude of the vertical compliances at the middle of the plate are calculated. The results shown in Fig. 2 indicate that the length of the model of a rigid track has to be at least 11.0 m in order to avoid large discretization errors. The amplitudes of the vertical compliances at the center of the concrete layer increase with a bigger length of the modelled plate. Convergence of the solution is seen by increasing the model length from 5 m to 11 m.

3.2   Influence of Different Track Geometry

Regarding the results of Sect. 3.1 by using a model length of 11.0 m, the vertical displacements of different rigid track systems on the half-space subjected to a harmonic load of $P = 0.5N$ at two rail support points as shown in Fig. 3 are investigated. The material properties, width b and hight h of concrete layer (CL), hydraulically consolidated layer (HCL), frost protection layer (FP) and the half-space are listed in Table 1.
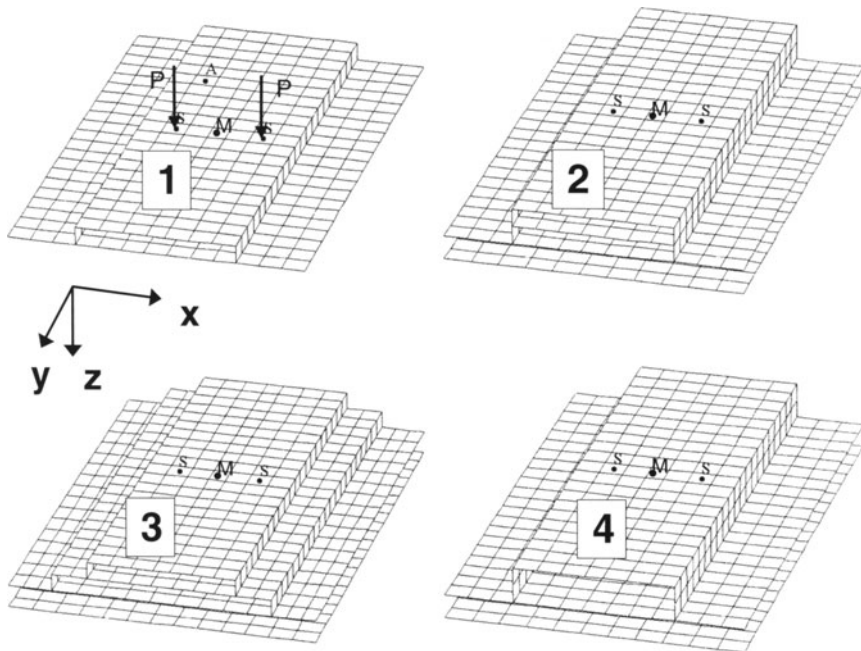
Fig. 3. Geometry of the rigid track systems 1 to 4

Figure 4 shows the vertical displacements of the analyzed rigid track systems in the middle of the concrete layer (point M). It can be seen that the

Table 1. Material properties and geometry of the models

	$c_s\,[m/s]$	$\varrho\,[kg/m^3]$	$\nu\,[-]$	$b\,[m]$	$h\,[m]$
System 1					
CL	2236	2500	0.2	2.7	0.3
Half-space	200	2000	0.33		
System 2					
CL	2407	2445	0.2	2.7	0.3
HCL	1903	2300	0.2	2.7	0.3
FPL	217	1938	0.32		0.4
Half-space	233	1837	0.33		
System 3					
CL	2407	2445	0.2	2.7	0.3
HCL	1903	2300	0.2	4.0	0.3
FPL	217	1938	0.32		0.4
Half-space	233	1837	0.33		
System 4					
CL	2236	2500	0.2	2.7	0.6
FPL	217	1938	0.32		0.4
Half-space	233	1837	0.33		
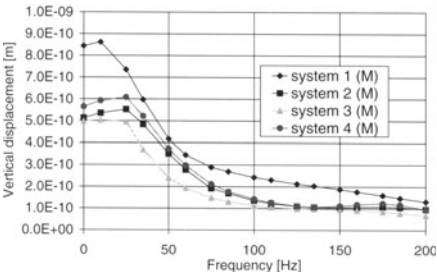
Fig. 4. Vertical displacements of systems 1 to 4, evaluated in point M
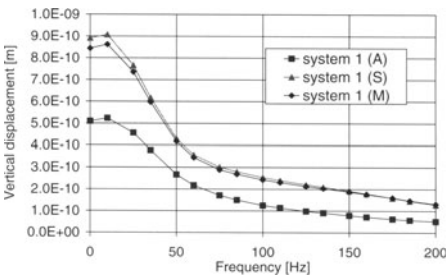
Fig. 5. Vertical displacements of system 1, evaluated in points S, M and A

displacements mostly depend on the material properties and thickness of the concrete layer. The amplitude characteristic shows that the main differences between the four systems become apparent in the frequency range between 0 Hz and 25 Hz. For frequencies higher than 25 Hz the displacements decrease monotonously with increasing frequency. A comparison of systems 2 and 3 which differ in the width of concrete layer and hydraulically consolidated layer, shows that the static displacements are almost identical while the bigger width of the hydraulically consolidated layer of system 3 leads to

a reduction of the displacements in the frequency range from 10 Hz up to 100 Hz. An investigation of the vertical displacements of system 1 evaluated in the middle of the concrete layer (point M), at the rail support points S and at a point A, which has a distance of 1.9 m from point S in y-direction, is shown in Fig. 5.

Owing to the high stiffness of the cross section of the concrete layer, only a small difference between the displacements in point M and point S can be observed. Similar results of an investigation of a rigid track are presented in [22], where an approach based on the coupling of the thin layer method with a finite element method is used. Measured results can be found in [20].

In addition to the results obtained by a load directly applied on the concrete layer, a further investigation is carried out on the influence of the railpads when the load acts on the rail. The displacements of the rail for a rigid track and a ballasted track on the top of an elastic halfspace as well as on the top of a rigid halfspace are given in Fig. 6 and Fig. 7. The results are compared with the displacements evaluated directly at the surface of the track at the rail support points below the load.
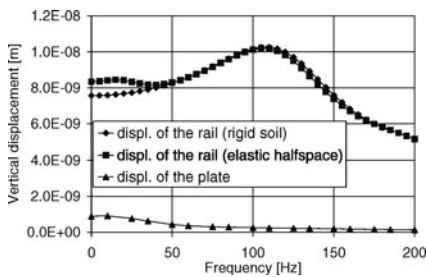
Fig. 6. Influence of the soil on the vertical displacements of a rigid track
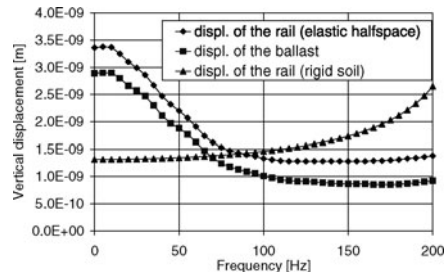
Fig. 7. Influence of the soil on the vertical displacements of a ballasted track

Due to the relatively small stiffness of the pads ($k = 2.2510^7\,N/m$) of a rigid track, the displacement at the rail is about 10 times higher than the displacement at the concrete layer of the rigid track (see Fig. 6). It can be seen, that only in the lower frequency range up to 40 Hz the displacements of the rail are determined by the properties of the subsoil. In the case of a ballasted track, as shown in Fig. 7, the displacements of the rail and the displacements of the ballast below the rail do not differ very much due to the fact that on the one side the ballasted layer is much softer than the rigid track and on the other side the stiffness of the railpad used for ballasted tracks ($k = 6.010^8\,N/m$) is higher than the elastic pad of a rigid track. In this case the numerical model for the ballast and the underlying subsoil is of great importance because the frequency dependency of the vertical displacements of a ballasted track show contrary behaviour in the case of a rigid or an elastic half-space (see Fig.

7). The calculation of the track-subsoil interaction considering the effect of vibration barriers is presented in [9]. A deeper investigation of the dynamic behavior of the railpads used for rigid tracks and ballasted tracks is carried out in [10] and [13]. An extended examination of the influence of the subsoil on the vertical displacements of the rigid track as well as of a ballasted track is presented in the frame of the benchmark test which is also published in this book.

3.3   Moving Load

Another goal of this project is the determination of the dynamic behavior of the track caused by loads moving with different velocities along the track. In a first investigation a pulse-type axle load, resulting from the distribution of the load from the contact point of the wheel over the rail, the elastic pads and the sleepers [27], [24] is directly applied to the surface of the concrete layer (see Fig. 8). In a second examination the load is moving along the rail supported by elastic pads and ties. Both investigations are based on the assumption that a point on the track subjected to a moving load experiences a temporal load of the same form as the spatially distributed load. The temporal load distribution P(t) used for both cases is shown in Fig. 8. Since the developed
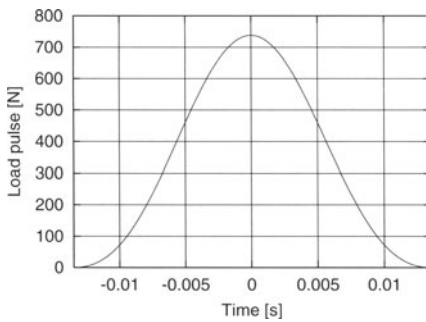
Fig. 8. Temporal load distribution for a one-axle load, v=400 km/h
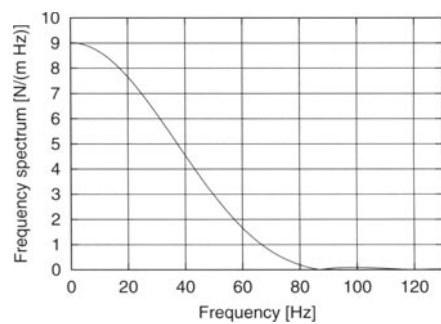
Fig. 9. Corresponding frequency spectrum $P_T(\omega)$

programm computes the responses in the frequency domain, the pulse load $P(t)$ has to be transformed via Discrete Fourier transform into frequency domain. The calculated displacements in the frequency-domain have to be transformed with the Discrete Inverse Fast Fourier Transform back into the time domain [4]. For a functions f and its transform F, the formulas for the Discrete Fourier Transform and the inverse Fourier Transform are given as:

$$F_N = \frac{T}{N} \sum_{k=0}^{N-1} f_k e^{-i2\pi \frac{nk}{N}}, \quad n = 0, 1, 2, ..., N-1 \tag{19}$$

and

$$f_k = \frac{1}{T} \sum_{n=0}^{N-1} F_n e^{i2\pi \frac{nk}{N}}, \quad n = 0, 1, 2, ..., N-1 , \tag{20}$$

where N is the number of the chosen time steps Δt and

$$T = N \, \Delta t \tag{21}$$

is the period. The frequency content $P_T(\omega)$ of the considered axle load is shown in Fig. 9. The displacements in the frequency domain $\boldsymbol{u}_T(\omega)$ are obtained from (18) with the load $P_T(\omega)$.
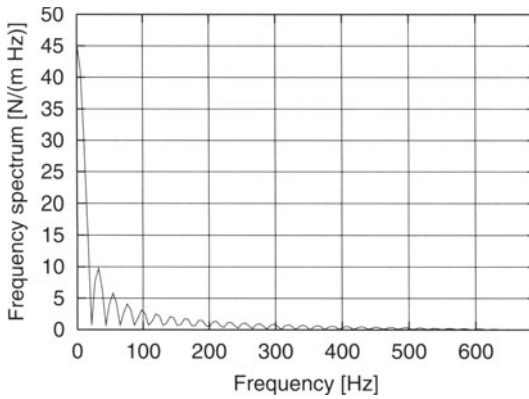
Fig. 10. Fourier Transform of the square-wave impulse

The implementation of the above mentioned algorithm is tested by the calculation of the displacement of two adjacent rigid ties at which one tie is subjected to a time dependent square-wave impulse of magnitude 1.0 kN whose frequency spectrum is shown in Fig. 10. The chosen time step is $\Delta t = 7.22^{-4}s$ and $N = 64$ is the number of time steps. The results are compared with a calculation performed by the Technical University Hamburg Harburg (TUHH) with their time-domain program. The results of the calculated displacements shown in Fig. 11. indicate good agreement between the two methods.

Other calculations in cooperation with the TUHH are presented in the paper of von Estorff at al. in this book.

Considering a moving load the time dependent response $u^k(t)$ at a selected point k on a track with length L is obtained by a superposition of the responses due to a load acting at discrete points i along the path. The time lag of the responses corresponds to the velocity of the moving load (see Fig. 12). The formula for the superposition is
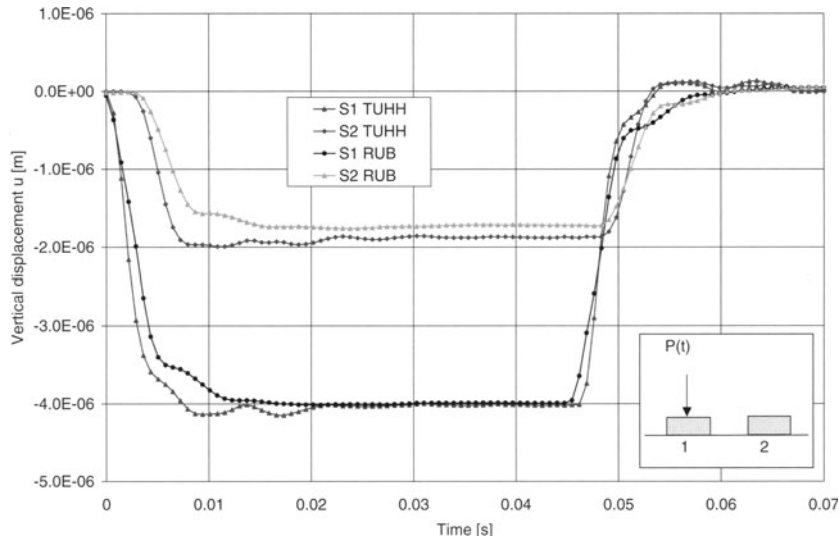
Fig. 11. Vertical time-dependent displacement of adjacent ties

$$u^k(n\,\Delta t) = \sum_{i=1}^{n} u^i((n-i+1)\Delta t) \; . \tag{22}$$

The distance Δe of two consecutive nodes depends on the velocity v of the
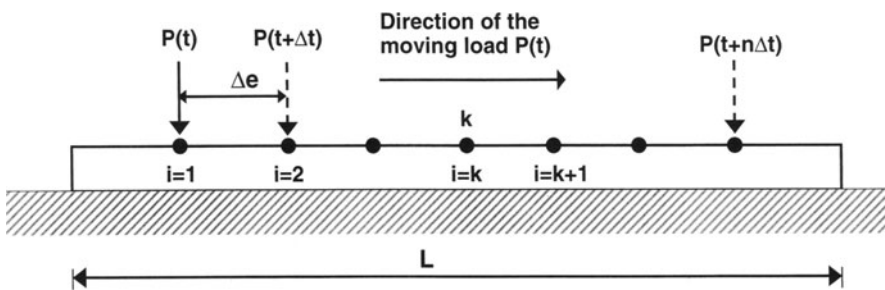
Fig. 12. Numerical calculation due to a moving load

load and the chosen time step Δt:

$$\Delta t = \frac{\Delta e}{v} \; . \tag{23}$$

The maximum frequency depends on the chosen time step given by

$$f_{max} = \frac{1}{2\Delta t} \,, \tag{24}$$

and the frequency step results from

$$\Delta f = \frac{1}{N\,\Delta t} \,. \tag{25}$$

In order to avoid long computing time, not every frequency in the required frequency range is computed. A well tested algorithm for a non-linear inter-polation of substeps which is based on the response of a 2-degree of freedom system [25] is implemented in the program. A comparison of the completely calculated frequency spectrum of a rigid track system with 41 frequencies and $\Delta f = 5\,Hz$ and an interpolated frequency spectrum based on only 15 calculated frequencies with $\Delta f = 15\,Hz$ is shown in Fig. 13. The calcu-lated frequency step of the interpolation procedure is $\Delta f = 1.6\,Hz$. Apart from a small difference in the lower frequency spectrum, the interpolated dis-placements show very good agreement with the results calculated with all 41 frequencies. Accordingly, the procedure is a suitable tool especially for a rigid track on the half-space whose dynamic behavior is approximately similar to a 2-degree of freedom system.
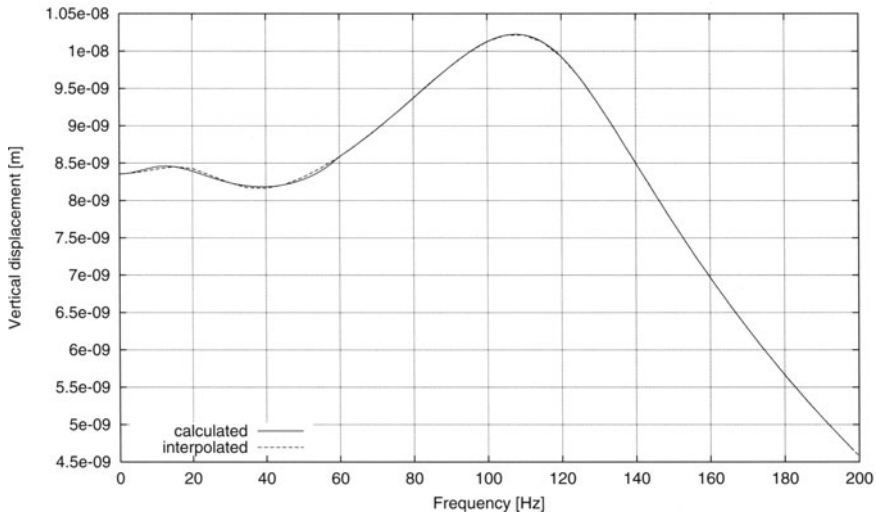
Fig. 13. Comparison of calculated and interpolated frequency spectrum

The algorithm presented in this chapter is applied to a rigid track resting on a half-space as shown in Fig. 14, where concrete layer, hydraulically consol-idated layer and frost protection layer are modelled in a simplified manner

as one layer with constant width. The material properties considered for the track and the varying soil are given in Table 2. The first calculation of the displacements of the concrete layer which is performed without rails, pads and ties is shown in Fig. 15, while the second calculation shown in Fig. 16 obtains the displacements at the rail due to the load acting along the rail.
In these calculations the pulse load results from a load velocity of 400 km/h (Fig. 8). The program allows the consideration of other pulse-type loads resulting from a superposition of the axle-loads of a moving train [26].
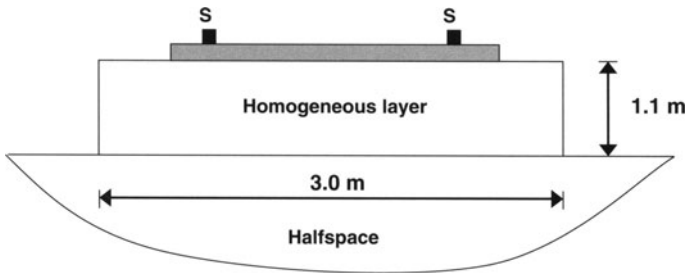
Fig. 14. Model of the rigid track

Table 2. Material properties and geometry of the model

	c_s [m/s]	c_p [m/s]	ϱ [kg/m³]
Rigid track	1070	1750	2150
Half-space 1	100	200	1700
Half-space 2	250	500	1700
Half-space 3	500	1000	1700

As shown in Section 3.2, the time history of the vertical displacements of the rail (Fig. 16), caused by a moving load are much bigger than the displacements of the concrete layer (Fig. 15). Due to the dominant influence of the elastic pads, the displacements caused by a moving load along the rail show no difference if half-space 2 or half-space 3 is considered, although half-space 2 is two times softer than half-space 3 as given in Table 2. Considering a moving load directly along the plate the various soils effect big differences in the vertical displacements of the concrete layer as shown in Fig. 17. More details on this calculation are given in [16] and [23].
Earlier results on the investigation of the influence of moving loads on a rigid track as well as directly on a half-space are presented in [24], [27], [7] and [26]. A semi-analytical algorithm for an investigation of a rigid track subjected to a harmonic moving load is shown in [5].
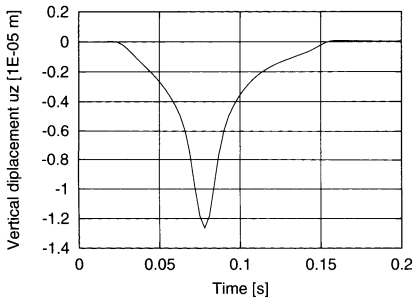
Fig. 15. Vertical displacement of the concrete layer; load on concrete layer
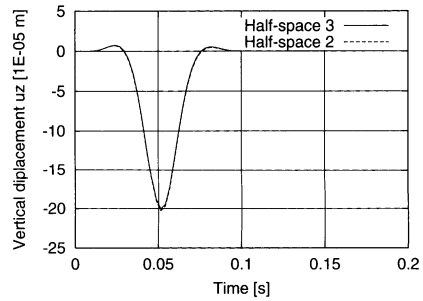
Fig. 16. Vertical displacement of the rail; load on rail
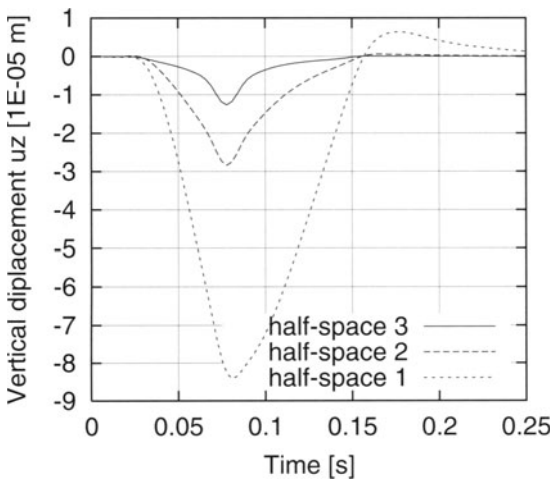
Fig. 17. Influence of different soils on the vertical displacements of the plate

4 Summary and Conclusion

The aim of this project is the development of an analysis procedure using the finite element method and the boundary element method to satisfy the various boundary and interface conditions of track, substructure and layered soil. The coupling of both methods proves to be an appropriate technique for the investigation of the dynamic behavior of the interaction between railway track systems and the subsoil. The analysis is performed with harmonic loads and moving loads. Moving loads are considered by a frequency-time-domain transformation of the system displacements due to load pulses resulting from the train passage. The presented results show the importance of a sufficient discretization of the track model. The investigation of different track systems shows that a simplified model of the subsoil is adequate for the calculation of

the dynamic behavior of a rigid track. Due to the stiffness of the superstructure the load is distributed along the plate and only small stresses are induced into the underlying soil. The most significant part of the system displacement is induced by the railpad. Relatively soft railpads cause high displacement amplitudes of the rails whereas at the concrete layer only 10 % of the rail displacement occur. Considering moving loads on a rigid track, the results show that the maximum displacements of the rails are mostly independent of the stiffness of the subsoil, whereas the dynamic behavior of the concrete layer resulting from a load moving directly along this layer is determined by the properties of the underlying soil. An extensive comparison of the presented results with other numerical procedures within a benchmark test shows the usefullness of the presented developments with respect to generality, flexibility and accuracy.

References

1. Becker, A. (1992) The boundary element method in engineering. McGraw-Hill, Berkshire
2. Bode, C. (2000) Numerisches Verfahren zur Berechnung von Baugrund-Bauwerk-Interaktionen im Zeitbereich mittels GEENscher Funktionen für den Halbraum. Veröffentlichungen des Grundbauinstitutes der Technischen Universität Berlin, Berlin 2000
3. Bode, C., Hirschauer, R., Savidis S.A. (2000) Three-dimensional time domain analysis of moving loads on railway tracks on layered soils. In: Chouw, N., Schmid, G. (Eds.) Wave 2000, Proc. of the International Workshop Wave 2000 Bochum, Germany, 13-15 December, 2000. Balkema, Rotterdam, 3-12
4. Brigham, E.O. (1995) FFT Schnelle Fourier-Transformation, 6. Auflage. R. Oldenburg Verlag, München, 1995
5. Dinkel, J. (2000) Ein semi-analytisches Modell zur dynamischen Berechnung des gekoppelten Systems Fahrzeug-Fahrweg-Untergrund für das Oberbausystem Feste Fahrbahn. Dissertation, Technische Universität München, August 2000
6. Dominguez J. (1993) Boundary elements in dynamics. Computational Mechanics Publications, Southhampton
7. Friedrich, K., Pflanz, G., Schmid, G. (1999) Modellierung des dynamischen Verhaltens der Festen Fahrbahn bei Zugüberfahrt. In: Feste Fahrbahn - Mechanische Modellierung, Betriebserfahrung und Akustik, 2. Auflage, Interdisziplinärer Forschungsverbund Bahntechnik IFV, Symposium 4.-5. Nov. 1999, Berlin
8. Grassie, S.L., Gregory, R.W., Harrison, D., Johnson, K.L. (1982) The dynamic response of railway track to high frequency vertical excitation. J Mech Eng Sci, Vol 24 No 2, 1982, 77-90

9. Hubert, W., Friedrich, K., Pflanz, G., Schmid, G. (2000) Frequency- and time-domain BEM analysis of rigid track and half-space with vibration barriers under investigation of causality errors for concave domains. EUROMECH Colloquium 414 Boundary Element Methods for Soil/Structure Interaction, Catania, Italy, Jun. 2000

10. Knothe, K. (2001) Gleisdynamik. Ernst & Sohn, Berlin

11. Lefeuve-Mesgouez, G. (1999) Propagation dondes dans un massif soumis à des charges se déplacant à vitesse constante. Thèse de Doctorat, Ecole Centrale de Nantes

12. Mohammadi, M., Karabalis, D.L. (1995) Dynamic 3-D soil-railway track interaction by BEM-FEM. Earthquake engineering and structural dynamics, Vol. 24, 1995. John Wiley & Sons, 1177-1193

13. Müller-Boruttau, F.H., Breitsamter, N. (2000) Elastische Elemente verringern die Fahrwegbeanspruchung. ETR Eisenbahntechnische Rundschau 49, Heft 9, Sept. 2001, 587-596

14. Petyt, M., Jones, C.J.C. (1999) Modelling of ground-bourne vibration from railways. In: Fryba, Naprstek (Eds.) Structural Dynamics, EURODYN 1999, Vol. I, Balkema, Rotterdam, 79-87

15. Pflanz, G. (2001) Numerische Untersuchung der elastischen Wellenausbreitung infolge bewegter Lasten mittels der Randelementmethode im Zeitbereich. VDI Fortschritt-Berichte Reihe 18, Nr. 265, VDI Verlag, Düsseldorf

16. Pflanz, G., Garcia, J., Schmid, G. (2000) Vibrations due to loads moving with sub-critical and super-critical velocities on rigid track. In: Chouw, N., Schmid, G. (Eds.) Wave 2000, Proc. of the International Workshop Wave 2000 Bochum, Germany, 13-15 December, 2000. Balkema, Rotterdam, 131-148

17. Przemieniecki, J.S. (1968) Theory of matrix structural analysis. McGraw-Hill, New York

18. Read, D., Matsumoto, N., Wakui, H. (1999) FAST testing Japanese-developed ladder sleeper system. RT&S: 16-18

19. Ripke, B. (1995) Hochfrequente Gleismodellierung und Simulation der Fahrzeug-Gleis-Dynamik unter Verwendung einer nichtlinearen Kontakt-mechanik. VDI Fortschritt-Berichte Reihe 12, Nr. 249, VDI Verlag, Düsseldorf

20. Rücker, W. (1999) Kurzzeitdynamik und Setzungsverhalten der Festen Fahrbahn. In: Feste Fahrbahn - Mechanische Modellierung, Betriebserfahrung und Akustik, 2. Auflage, Interdisziplinärer Forschungsverbund Bahntechnik IFV, Symposium 4.-5. Nov. 1999, Berlin

21. Rücker, W. (1981) Dynamische Wechselwirkung eines Schienen-Schwellensystems mit dem Untergrund. Forschungsbericht 78, Bundesanstalt für Materialprüfung, Berlin

22. Savidis, S.A., Bergmann, S., Bode, C., Hirschauer, R. (1999) Dynamische Wechselwirkung zwischen der Festen Fahrbahn und dem geschichteten Untergrund. In: Feste Fahrbahn - Mechanische Modellierung, Betriebserfahrung und Akustik, 2. Auflage, Interdisziplinärer Forschungsverbund Bahntechnik IFV, Symposium 4.-5. Nov. 1999, Berlin

23. Schmid, G., Pflanz, G., Friedrich, K., Garcia, J., Hubert, W. (2001) Dynamic soil-structure interaction for moving load - methods and numerical application. Proc. of 2nd ICTACEM at Kharagpur, India, Dec. 27-30, 2001. Indian Institute of Technology, Paper No. 113

24. Schmid, G., Verbic, B. (1997) Modellierung der Erschütterung aus dem Schienenverkehr mit der Randelementmethode. In: Bachmann, H. (Ed.) Erdbebensicherung bestehender Bauwerke und aktuelle Fragen der Baudynamik, Tagungsband D-A-CH'97, SIA, Dokumentation DO145, 1997

25. Tajirian, F. (1981) Impedance matrices and interpolation techniques for 3-D interaction analysis by the flexible volume method. PhD. Thesis, University of California, Berkeley

26. Tosecký, A. (2001) Numerische Untersuchung der Erschütterungsausbreitung infolge bewegter Lasten auf einem Feste Fahrbahn-System mittels der Methode der dünnen Schichten/Methode der flexiblen Volumen. Diplomarbeit, Ruhr-Universität Bochum, Bochum 2001

27. Verbic, B., Schmid, G., Köpper, H.D. (1997) Anwendung der Randelementmethode zur dynamischen Berechnung der Festen Fahrbahn, EI - Der Eisenbahningenieur, Februar 1997

Experimental and Numerical Investigations on the Track Stability

Rolf Katzenbach[1] and Stefan A. Heineke[2]

[1] Technische Universität Darmstadt, Institute and Laboratory of Geotechnics, Petersenstr. 13, 64287 Darmstadt, Germany
[2] IBES Baugrundinstitut GmbH, Neustadt/W., Germany
  formerly: Technische Universität Darmstadt, Institute and Laboratory of Geotechnics, Petersenstr. 13, 64287 Darmstadt, Germany

Abstract. In the paper a basic description of the settlement behaviour of a track due to the passing of trains is presented. Based on that fundamental consideration model scale tests were performed. As a result of these model scale tests a mathematical framework to predict the mean settlements and the differential settlements are presented. The transfer of the results of the model scale tests to a prototype is presented. In a numeric study the suitability of elasto-plastic material laws to simulate the settlement behaviour of a foundation under cyclic load is investigated. Furthermore a study is made to find out, if it is possible to simulate the moving of non-uniform settlements under a moving load in a numeric simulation.

1 Introduction and Motivation

The settlement behaviour of a track is in the geotechnical view characterised by an intensive soil-railway-carriage-interaction. The mechanical loading can be divided in a short-term dynamic loading and in a cyclic static loading. Between the short-term dynamic loading and the long-term behaviour of the track stability is a close feed-back. Non-uniform settlements leads to irregularities of the track, which has negative effects among other things on the following components:

- carriage
- travelling comfort
- safety
- travelling time
- maintenance
- susceptibility to trouble
- availability.

Within the project "Model scale investigations to the long-term behaviour of the track under railway traffic" the long-term behaviour of a track on a granular subsoil should be investigated. The aim of this project was, to find a mathematical description to predict the long-term settlement behaviour. Therefore a testing device was planned and several model scale tests were performed.

2   Problem

Irregularities of the track occurs at the passage from earthwork-lines to structures due to different stiffnesses of the underground and on the free line section as result of irregularities in the underground, which means that there is a difference in the first settlement and of course of modifications of the material of the granular soil, i.e. abrasion. To investigate the settlement behaviour of a track on a granular subsoil, a difference is made between uniform settlements and non-uniform settlements. With non-uniform settlements there is an appearance of waves. If waves appear it is of interest, wether the waves move in the driving direction of the running train. The aim is, to registrate the settlements in a geotechnical model scale test and to transfer the results on a prototype.

3   Solving Method

For the preformance of the model scale tests, the following assumptions were taken:

- A continous moved load has an influence on the initiation of "waves". Therefore a circular testing container is choosen in which the load is all the time running on the track and where it is not moved on or removed from the track for each passage.
- The minimal circumference of the testing container must be fixed in a way, that a settlement mould can be developed under the load. Therefore pre-tests were performed to obtain the maximal length of a settlement mould under different elastic slabs (Katzenbach et al., 1998).
- Investigations were made for a short train, i.e. the influence of the distribution of the load is neglected. Therefore the distributed load is reduced to a single load.
- Long intervals between the trains, i.e. successive trains do not lead to a superposition of the loads. Fixation of the velocity in a way, that effects of inertia can be excluded, so the quasi-static effects of the railway traffic are represented.

Fig. 1 shows the simplified model.
For the simplified model the settlements can be described by the following equation

$$s = f_1(x, N, P, b, EI/b, \gamma, K_0, \mathcal{G}) \tag{1}$$

with s settlement, x position, N number of loading, P load of the train, b width of the track, EI/b bending resistance over the width of the track, γ unit weight of the soil at the beginning of the experiment, K_0 virgin configuration of the
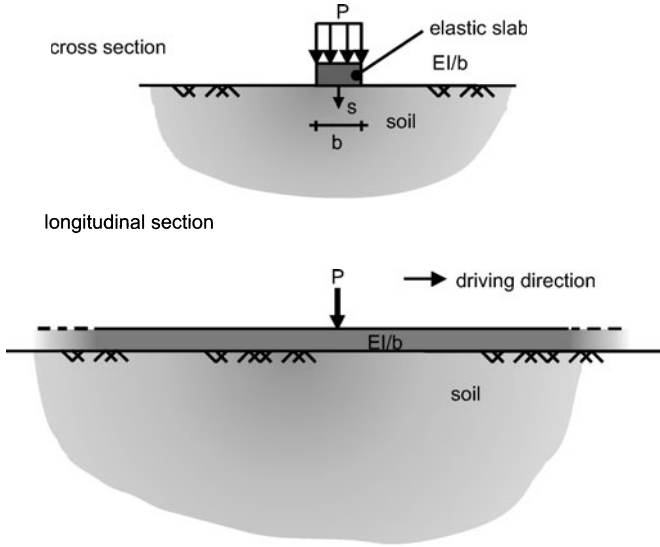
Fig. 1. Sketch of the simplified model

soil expressed by the coefficient of earth pressure at rest, $\mathcal{G}$ granulometry of the soil.

The model laws used in the investigations are based on Dietrich (1977, 1979, 1981, 1982). For the stress-strain-relation of the soil the following power function is used:

$$\Delta\varepsilon \sim \left(\frac{\Delta\sigma}{\gamma \cdot z}\right)^{\mu} \tag{2}$$

with ε strain, σ stress, z depth, μ material parameter.

The elastic slab is described by its curvature:

$$\kappa(x,t) = \frac{m(x,t)}{\frac{EI}{b}} \tag{3}$$

with κ curvature, m bending moment related to the width

With the deformation modulus of the soil D_U

$$D_U = \frac{1}{\gamma^{\mu}} = \gamma^{-\mu} \tag{4}$$

and the deformation modulus of the elastic slab D_O

$$D_O = \frac{1}{\frac{EI}{b}} = \left(\frac{EI}{b}\right)^{-1} \tag{5}$$

the rigidity of the system is defined as follows

$$\frac{D_U}{D_O} := \frac{EI}{\gamma^\mu \cdot b} \tag{6}$$

The rigidity of the elastic slab and the unit weigth of the soil at the beginning of the test is replaced an elastic length l_E which is defined as follows:

$$l_E := \sqrt[4]{\frac{EI}{\gamma \cdot b}} \tag{7}$$

The load of the train is replaced by the system length l_P which is defined as follows:

$$l_P := \left(\frac{P}{\gamma \cdot l_E^2}\right)^{\frac{\mu-1}{2(\mu+1)}} l_E \tag{8}$$

With the elastic length (7) and the system length (8) the rigidity of the system (6) becomes

$$s = f\left(l_P, \frac{D_U}{D_O}, D_O, \Pi\right) \text{ with } \Pi = \left(\frac{x}{l_P}, N, \frac{b}{l_P}, K_O, \mathcal{G}\right) \tag{9}$$

So the settlement is proportional to the system length, the rigidity of the system, the deformation modulus of the elastic slab and the function P as shown in Eq.10.

$$s \sim l_P \cdot \frac{D_U}{D_O} \cdot D_O \cdot \Pi \tag{10}$$

With aid of the dimension analysis the dimensionless description of the settlements of an elastic slab becomes:

$$\frac{s}{l_E} = \left(\frac{P}{\gamma \cdot l_E^2}\right)^{\frac{5\mu-1}{2(\mu+1)}} \Psi\left(\frac{x}{l_P}, N, \frac{b}{l_P}, K_0, \mathcal{G}\right) \tag{11}$$

To classificate a model system, Dietrich introduces a dimensionless parameter called "width ratio", which is defined as the width of the elastic slab divided by the system length as shown in Eq. 12.

$$\beta_P = \frac{b}{l_P} \tag{12}$$

With the width ratio a model system can be classified in a system with only uniform settlements ("rigid track") and in a system with non-uniform settlements ("elastic track"). Between the two extremes there is a transitional stage. The area limits of the width ratio are shown in Fig. 4.
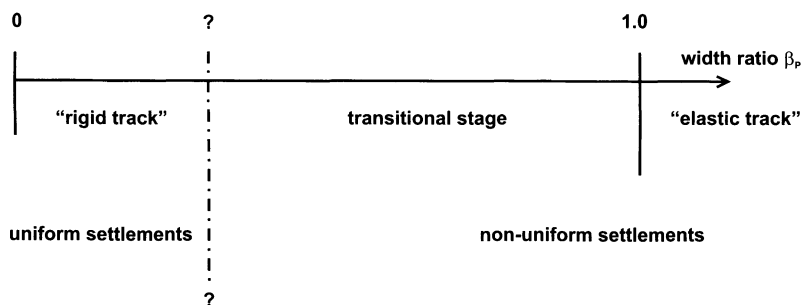
Fig. 2. Area limits of the width ratio
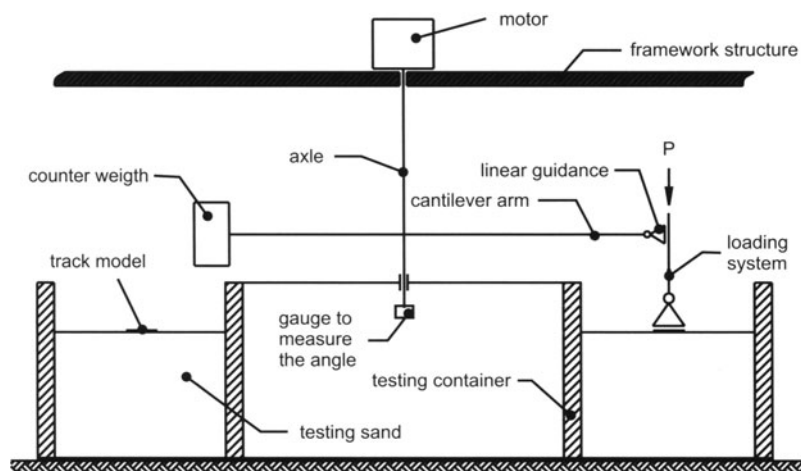
Fig. 3. Scetch of the testing stage

4   Model Scale Tests

The model scale tests were performed in a circular testing stage. A scetch of the testing stage is shown in Fig. 2.

The testing container for the determination of the long-term behaviour of the track due to plastic deformations of the subsoil was planned as a circular testing container with an outer diameter of 2.0 m and an inner diameter of 1.18 m. The average circumference measures 5 m. The testing container consists of two radial bricks of concrete. The surface was treated to obtain a smooth surface. The deformation is measured by a dial gauge in combination with a gauge to measure the angle. Fig. 3 shows a photo of the circular testing device and a detail of the loading system.

For the model scale tests a dry sand with a coefficient of uniformity of U=3 was used. The grain size distribution of the testing sand is given in Fig. 5. The filling and preparation of the testing container is described in Heineke (2001).
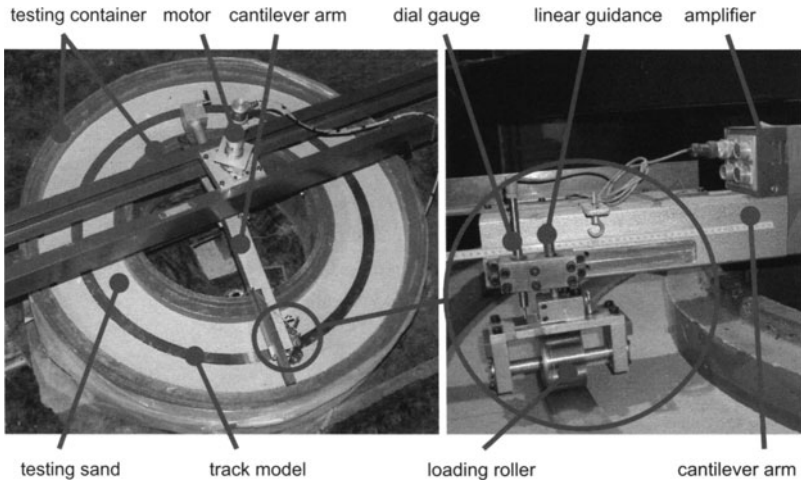
Fig. 4. Photo of the testing device and detail of the loading system
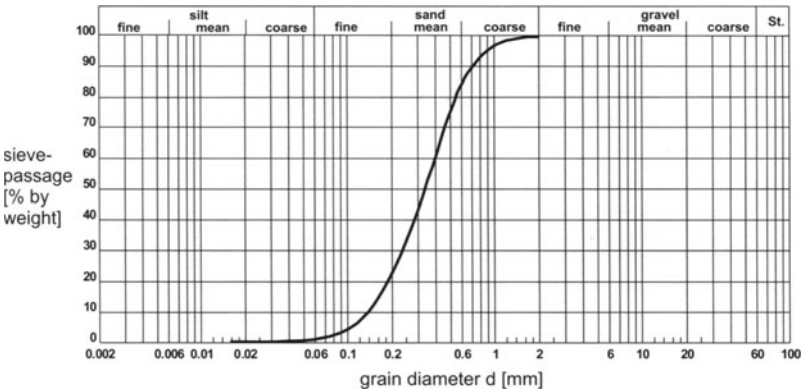
Fig. 5. Grain size distribution of the testing sand

The following figure shows the parameter of the tests presented in this paper. The velocity of the loading roller was chosen to 10 rounds per hour. The result is, that there are no effects of inertia, but that one test with 50000 loadings takes about 7 months.

The model scale tests showed, that on an initial homogeneous soil non-uniform settlements can occur with the passing of a moving load. With each passage of the load the settlements increase. Fig. 7 shows exemplary the development of the settlements over the position of the load in the testing container for the model scale test HV3. The curve parameter represents the number of loading.

test / parameter	HV_3	HV_8	HV_10	HV_12
symbol	▲	■	●	◆
width b [mm]	60	60	40	60
height h [mm]	5	15	15	15
unit weight of the soil at the beginning of the test γ [kN/m³]	17	17.5	17	16.8
load [N]	30	30	30	30
material of the elastic slab	polypro-pylen	polypro-pylen	aluminium	aluminium
E-modulus of the elastic slab [MN/m²]	1300	1300	70000	70000
μ [-]	1.97	1.97	1.97	1.97
elastic length l_E [m]	0.168	0.380	1.037	1.040
system length l_P [m]	0.169	0.292	0.616	0.579
width ratio β_P [-]	0.355	0.206	0.065	0.103

Fig. 6. Overview of the testing parameter
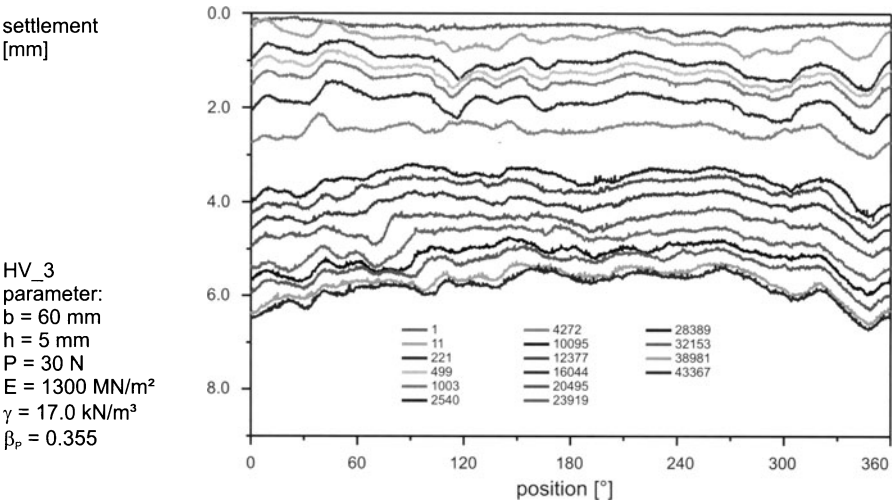
Fig. 7. Development of the settlements for the model scale test HV3

In the model scale tests the moving of settlement-moulds in the direction of the moving load has also be recognised. Some more information can be found in Heineke & Katzenbach (2001).

Figure 8 shows a comparison between the settlements of the different tests mentioned in Fig. 6. To compare the tests the settlements are related to the system length to have a dimensionless parameter. The curves are plotted for about 50000 number of loadings. It can be seen, that the settlements becomes greater and more non-uniform with an increasing width ratio, that means with softer systems. With the investigations the area of the rigid track,

that means that there are uniform settlements, has been recognised within some tolerances.
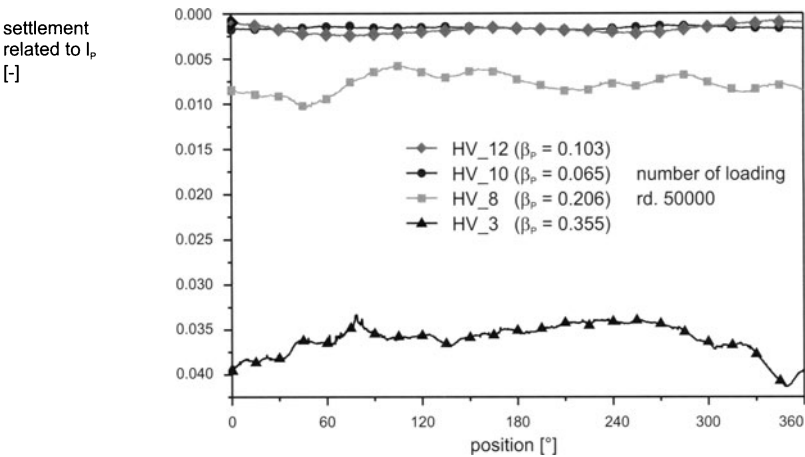
Fig. 8. Related curve of the settlements for different model scale tests

The related mean settlements of the track in dependence of the number of loading is shown in Fig. 9. Here as well it can be seen, that the mean settlements becomes greater with an increasing width ratio but that the increment of the settlements becomes smaller with increasing number of loadings and it seems that they reach a limiting value.
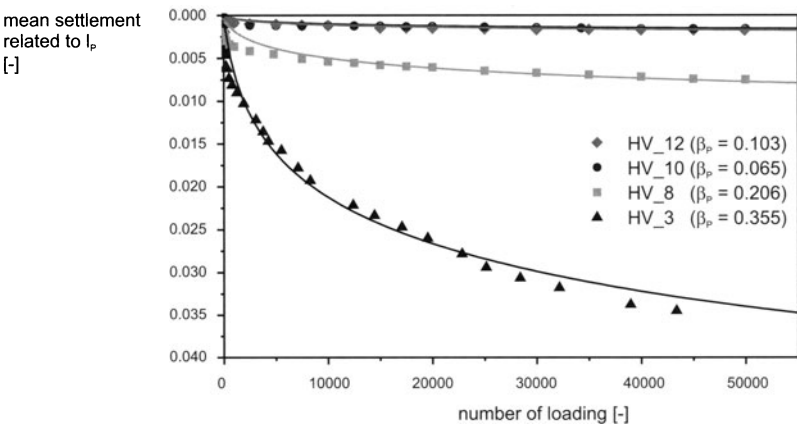
Fig. 9. Related mean settlement of the tack in dependence of the number of loading

The development of the mean settlements can be described for each model scale test by the following new logarithmic function.

$$\frac{s_M(N)}{l_P} = \frac{s_M(N=1)}{l_P} + a \ln \left(\frac{N-1}{1000} + 1 \right) \tag{13}$$

with s_M (N) mean settlement of the track at the number of loading N, a constant of the test, to be determined out of the model scale test

The plotted lines in the graph of Fig. 9 are representing this function. A good description of the measured settlements can be determined with Eq. (13). If the first settlement and the constant a is known, it is possible to predict the development of the settlements for different systems. Therefore a relation between these parameters and the width ratio was found. Figure 10 shows the relation between the first settlement related to the system length and the width ratio. It can be seen, that there is a nearly linear dependence between the two parameters, which can be described by the linear function of Eq. (14). The relation between the constant a and the width ratio, shown in Fig. 10 can be described by the non-linear function of Eq. (15).
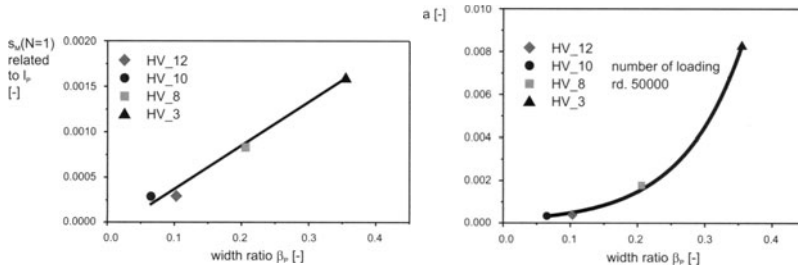
Fig. 10. Development of the mean first settlement (left) and the constant a (right) in relation to the witdh ratio

$$\frac{s_M(N=1)}{l_P} = 0.004789 \cdot \beta_P - 0.00011 \tag{14}$$

$$a = 0.00016 \cdot e^{11.12 \cdot \beta_P} \tag{15}$$

With the knowledge of these two parameters, that are in relation with the width ratio, it is possible to predict the mean settlement for different track systems on a granular soil.

In addition to the knowledge of the mean settlement it is of importance to know the differential settlements, to be able to make a prediction of the non-uniform settlements. Therefore a nearly linear relation between the differential settlement at about 50000 loadings and the width ratio was found. The relation between the two parameters is shown in the graph of Fig. 11 and

the mathematical description in Eq. (16). The number of loading of 50000 is sufficient, because the waviness or in a more mathematical description the standard deviation from the mean settlements becomes constant (Heineke et al., 2001).
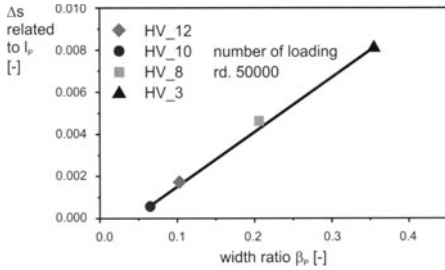
Fig. 11. Development of the differential settlement in relation to the width ratio

$$\frac{\Delta s}{l_P} = 0.026 \cdot \beta_P - 0.00111 \tag{16}$$

5   Transfer of the Results on a Prototype

An example to transfer the obtained results of the model scale tests on a prototype is given for a slab track "system Züblin". The slab track is simplified in the following components (see also Fig. 12):

- rail UIC 60
- concrete base (BTS)
- stabilised soil base (HGT)
- soil
- load from an ICE.

As an assumption a transferable power between the rails, the concrete base and the stabilised soil base is given. With this assumption a bending resistance of EI=990 MN/m^2 is obtained. The elastic length Eq. (7) is calculated to 11.66 m and the system length Eq. (8) to 8.22 m. With this input the width ratio Eq. (12) is obtained to 0.383. The width ratio is now the input parameter to calculate the settlement after the first loading, the constant a and the differential settlement as follows:

$$\frac{s_M(N=1)}{l_P} = 0.004789 \cdot \beta_P - 0.00111 = 0.04615 \tag{17}$$

$$a = 0.00016e^{11.12 \cdot \beta_P} = 0.01130 \tag{18}$$
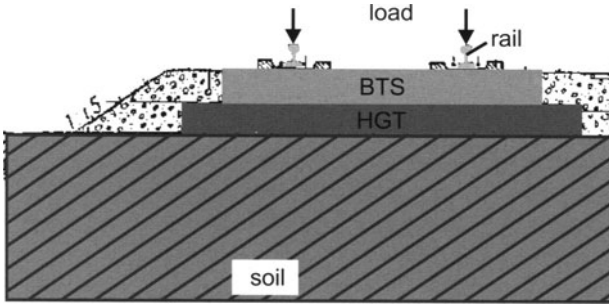
Fig. 12. Simplified slab track of the "system Züblin"

$$\frac{\Delta s}{l_P} = 0.026 \cdot \beta_P - 0.00111 = 0.008848 \tag{19}$$

With Eq. (17) and Eq. (18) in Eq. (12) the settlements of the prototype are obtained to:

$$s = \left(\frac{P}{\gamma \cdot l_E^2}\right)^{\frac{5\mu-1}{2(\mu+1)}} \cdot l_E \cdot \frac{s_M(N=1)}{l_P} + a \ln\left(\frac{N-1}{1000} + 1\right) = 2.3cm \tag{20}$$

The differential settlement is obtained with Eq. (19) and Eq. (12) to:

$$\Delta s = \left(\frac{P}{\gamma \cdot l_E^2}\right)^{\frac{5\mu-1}{2(\mu+1)}} \cdot l_E \cdot \frac{\Delta s}{l_P} = 0.44cm \tag{21}$$

For the slab track "system Züblin" the mean settlements are calculated to 2.3 cm and the differential settlement to 0.44 cm. So the differential settlement is about 20 % of the mean settlement.

6   Numeric Simulation

In the numeric simulation a study was made with a foundation under cyclic loading to investigate the suitability of elasto-plastic material laws. The results of this investigation is documented in Heineke (2001).

Furthermore the pre-test was simulated on the basis of the material law of von Mises with kinematic hardening in a finite element model. Figure 13 shows the measured settlements in the pre-test. In the graphic the settlements are plotted over the distance from the loading point. The curve parameter represents the number of loading.

Figure 14 shows the numeric model of the pre-testing stage to simulate the measured settlement-moulds of the pre-test.

The soil was simulated with continuum elements (4-nodes) and the elastic slab by beam elements (2-nodes). The cyclic loading was simulated in the
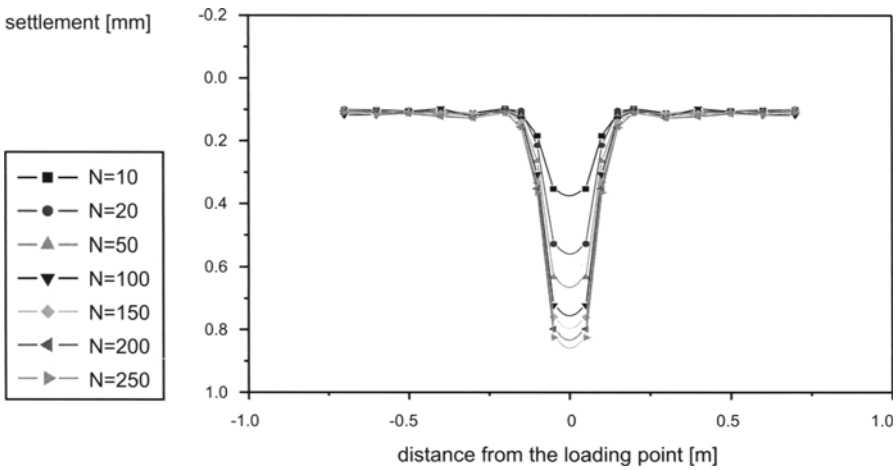
Fig. 13. Measured settlement-mould under the cyclic loading in the pre-test
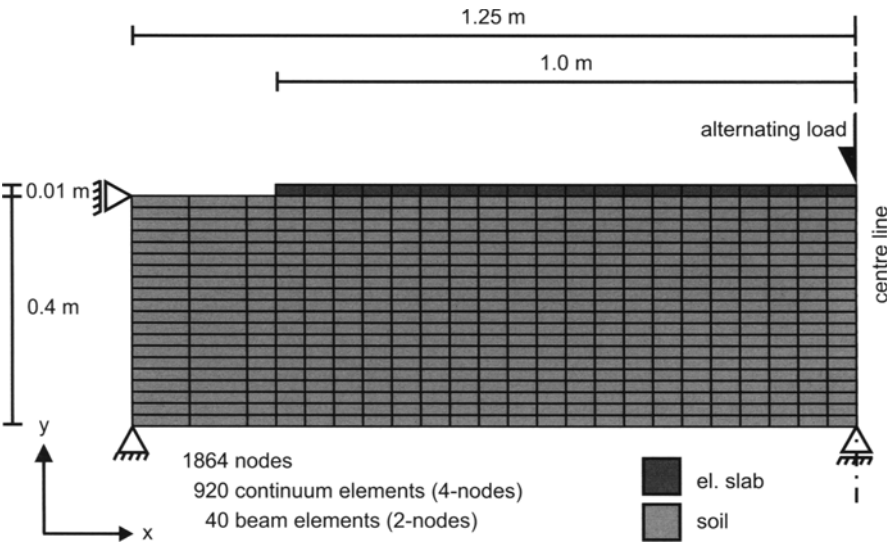
Fig. 14. Numeric model of the pre-testing stage

center line of the system in the middle of the elastic slab. The geometric dimension is given in Fig. 14.

Figure 15 shows the calculated settlement-moulds of the numeric simulation. It can be seen, that there is a relativ good conformity between the measured and the calculated settlement-moulds for a small number of loading (Fig. 15), but for high number of loadings the conformity is not good at all.

In a second model a study was made, wether it is possible to simulate the moving of the settlement-mould under a moving load in one direction. Therefore
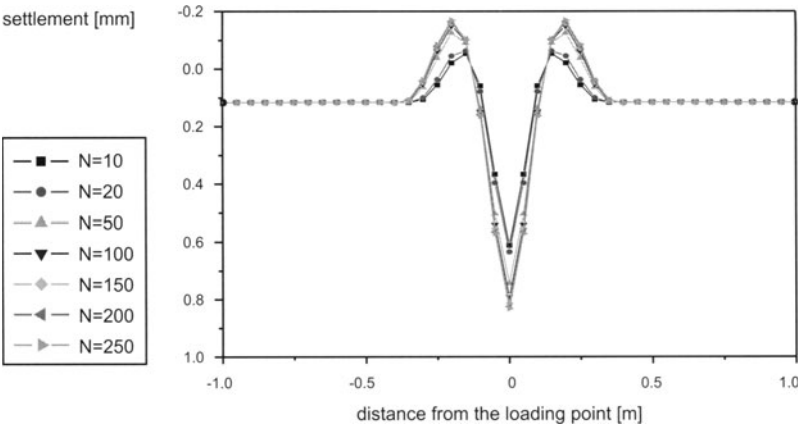
Fig. 15. Numeric model to calculate the moving of waves

the soil and the elastic slab was simulated by continuum elements (8-nodes). The load was simulated as a moving load in one direction. Figure 16 shows the numerical model to calculate the moving of waves.
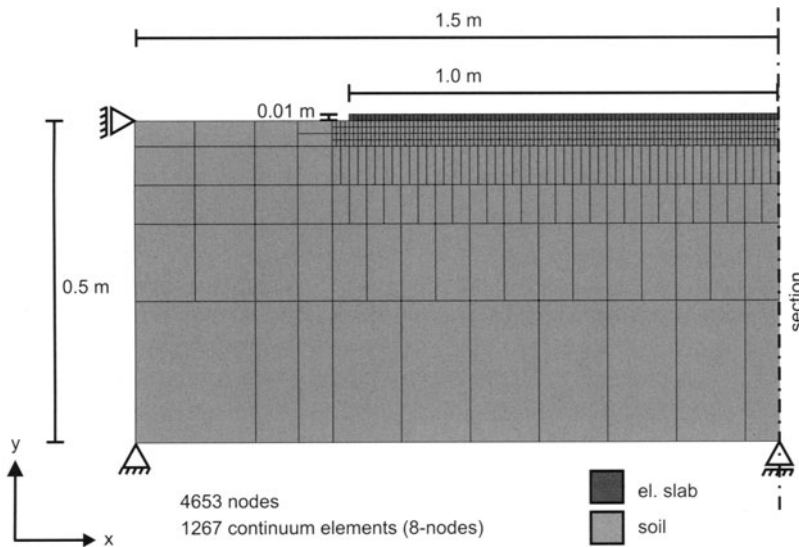
Fig. 16. Numeric model to calculate the moving of waves

Figure 17, 18 and 19 shows the settlement moulds of three different systems. The rigidity of system 2 is about two times bigger than system 1 and system 3 is three times bigger than system 1. The interpretation of the simulations show, that the settlement-mould becomes longer with a rigid elastic slab

and that the maximal settlement becomes smaller of course of the extended stress under the slab on a bigger area, what is documented by a longer elastic length. Furthermore the settlement mould is moving in the direction of the load. The moving of the settlement-mould is as well inclined by the rigidity of the system.
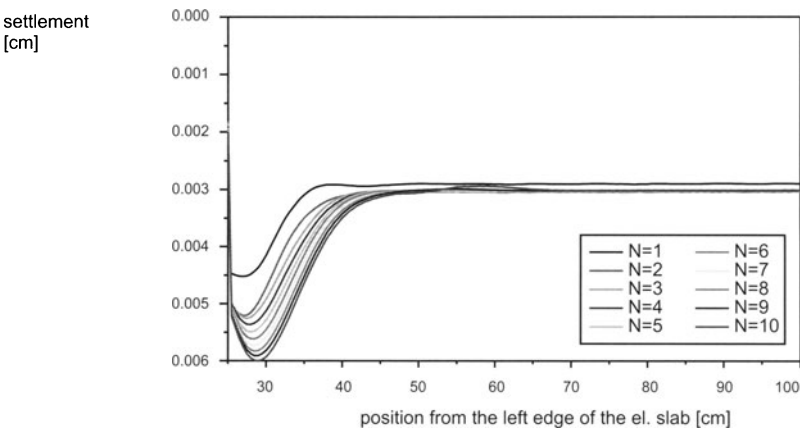
Fig. 17. Development of the settlement-mould for system 1
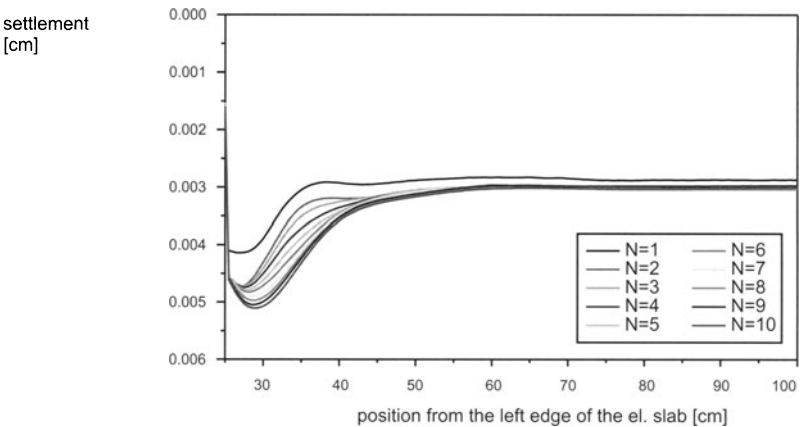
Fig. 18. Development of the settlement-mould for system 2

The numeric investigation showed, that it is possible to simulate the moving of waves in a qualitativ way for some passings of the load but that the use of the conventional elasto-plastic laws is not really good to simulate a system under cyclic load.
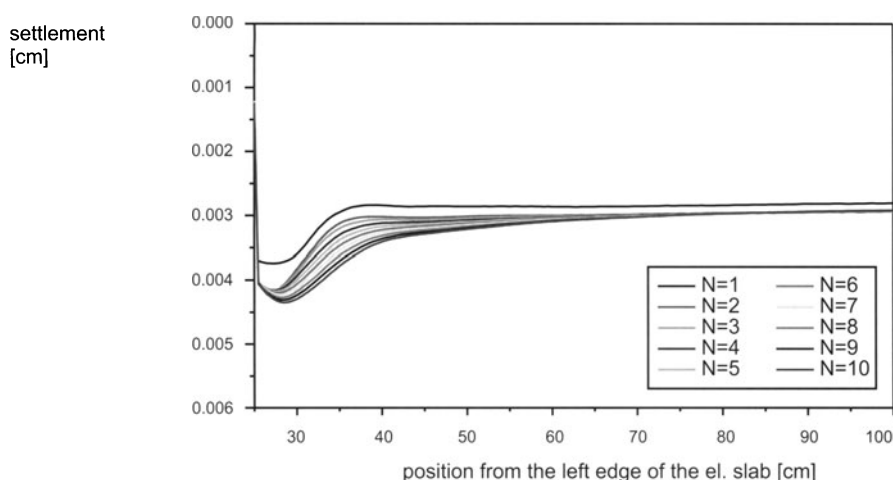
Fig. 19. Development of the settlement-mould for system 3

7 Possible Application of the Research Results to the Praxis

Up to now the dimensioning of slab tracks is based on a proof of the permitted tensile strength in bending, which is a inner dimensioning of the system. The settlements respective the differential settlements were not considered. A possible application of the research results to the praxis is, that the settlements or differential settlements were considered (long-term behaviour) in a proof of the serviceability of the system. In combination with the classic dimensioning an inner and outer dimensioning of the system should be established.

8 Summary

Summing up for the experimental investigations it can be noticed:

- the mean settlements grow with increasing number of loading
- the strain increment becomes smaller with increasing number of loading
- discovery of a new empiric relation to obtain the mean settlements
- the differential settlement grows with increasing number of loading and approximates to a limiting value
- the differential settlement is described by a new empiric relation
- functional relations under consideration of the width ratio has been found
- the results of the model scale tests can be transmitted with the similarity mechanics on prototypes.

Summing up for the numerical investigations it can be noticed:

- with classic elasto-plastic laws a monoton cyclic load can only be simulated for a small number of loading
- for a great number of loading the plastic increment becomes zero
- with kinematic elasto-plastic laws a constant plastic incremant is obtained
- the movement of waves can be simulated numerically, but this takes a lot of time with the actual capacity of personal computers.

9   Outlook

In the future the following investigations and research activities can be made:

- adaption of the obtained results to a mechanic model
- verification to transfer the obtained results on a cohesive soil
- development of a new elasto-plastic laws to simulate cyclic loads and of a qualified interaction model
- execution of numeric simulations and calibration with the performed model scale tests

References

1. Katzenbach R., Arslan U., Heineke St. (1998) Modellversuche zu den Langzeitveränderungen der Gleislage infolge plastischer Deformationen des Untergrundes. DFG-Schwerpunktprogramm "Systemdynamik und Langzeitverhalten von Fahrwerk, Gleis und Untergrund", Arbeitsbericht vom 24.3.1998
2. Dietrich Th. (1977) Der psammische Stoff als mechanisches Modell des Sandes. Dissertation, Institut für Bodenmechanik und Felsmechanik, Universität Fridericiana Karlsruhe
3. Dietrich Th. (1979) Ermittlung der plastischen Verformungen des Untergrundes infolge Schienenverkehr und Shakedownverhalten rolliger Böden für den Grenzfall starrer Körner. 1. und 2. Technischer Bericht zu TV 78172 Teil C, Bundesanstalt für Materialprüfung, Berlin
4. Dietrich Th. (1981) On the validity of Winkler's principle. Proc. Xth ICSMFE **2**, 103-108
5. Dietrich Th. (1982) Ermittlung der plastischen Verformungen des Untergrundes infolge Schienenverkehr und Shakedownverhalten rolliger Böden für den Grenzfall starrer Körner. 5. Technischer Bericht zu TV 78172 Teil C, Bundesanstalt für Materialprüfung, Berlin
6. Heineke St. (2001) Experimentelle und numerische Untersuchungen zur Gleislagestabilität. Mitteilungen des Institutes und der Versuchsanstalt für Geotechnik **57**, Technische Universität Darmstadt
7. Heineke St., Katzenbach R. (2001) Gebrauchstauglichkeitsuntersuchungen zur Festen Fahrbahn im Modellversuch. Bauingenieur 76 **6**, 276-282
8. Heineke St., Katzenbach R., Arslan U. (2001) Model scale investigations on the deformation of the subsoil under railway traffic. Proc. XVth ICSMGE **3**, 2077-2080

Experimental Investigation and Numerical Modelling of Soils and Ballast under Cyclic and Dynamic Loading

Hans-Georg Kempfert[1], Alexander Gotschol[1], Yifeng Hu[2], and Tim Stöcker[1]

[1] Kassel University, Fachgebiet Geotechnik,
   Mönchebergstr. 17, 34127 Kassel, Germany
[2] LGA, 90431 Nürnberg, Germany

Abstract. It is reported on the knowledge advance, regarding the long-term deformation behaviour of soils under cyclic-dynamic loading. A constitutive approach for water-saturated soft soils and its numerical implementation is presented. Furthermore the modification and enhancement of the constitutive approach onto granular materials and ballast by means of a cyclic-viscoplastic approach is described. In addition, it is reported on the implementing of the new cyclic-viscoplastic constitutive approach into the FE code ABAQUS. Hence, in dependence of the essential loading condition and material parameters, the description of the plastic long-term behaviour under cyclic-dynamic loading is made possible. Some verification and calculation results are presented which demonstrate the modelling capabilities of the implemented cyclic-viscoplastic approach.

1 Introduction

This report describes the final and preceding stages of the "System Dynamics and Long-term Behaviour of Vehicle, Track and Subgrade (DFG Priority Program)" project. During the first and second phase (05/96 - 04/00) a substantial advance has been gained regarding the long-term behaviour of fully saturated soft soils under cyclic loading. A quasi-static constitutive approach for such material has been derived and implemented into a numerical program. Within the third phase of the priority program, the modification, i.e. the extension of the research advances gained so far, onto granular materials and ballast is carried out. Beyond the long-term deformation behaviour, the altering (short-term) dynamic material properties are investigated. A combined cyclic triaxial resonant column test device (CTRC) that has been developed beforehand phase three, is used extensively. Next to the investigation of real scale ballast materials in cyclic triaxial tests, the CTRC makes possible the identification of altering short-term parameters vs. loading time/load-cycles, such as shear modulus and damping coefficient. Hence, constitutive derivatives for the mentioned short-term material parameters that should be used for dynamic modelling of the structure are acquired. Thereupon a newly derived cyclic-viscoplastic constitutive approach for granular soils and ballast

materials is derived and implemented into a quasi-static numerical model for the simulation of the long-term behaviour of such materials.

2   Saturated Cohesive Soils

2.1   General

This section focus on the plastic strain of normally to lightly overconsolidated cohesive soils under cyclic loading. Excess pore water pressure rises in the first stage and consists of reversible and "permanent" components $(u = u^c + u^b)$. The "permanent" part u^b results from the undrained plastic shear deformation, and remains when load is removed but dissipation not allowed. In the course of cyclic loading, "permanent" excess pore water pressure reaches a peak value at some time t_p and begins to dissipate if possible (the second stage), which leads to plastic volume strain. The peak of excess pore water pressure and t_p are dependent on e.g. cyclic loading, static preloading, permeability, drainage path and boundary conditions. Compared to the "permanent" part, the reversible part of excess pore water pressure as well as deformation is normally much smaller. Furthermore, this part of excess pore water pressure can not dissipate due to its oscillating character. In the following, a quasi-static model and its numerical implementation is described for investigating the development of excess pore water pressure and plastic strain of normally consolidated or lightly overconsolidated cohesive soils under cyclic loading, see [1]. It is verified by using test results from literature. Some own results of cyclic triaxial tests on a medium plasticity clay are presented.

2.2   Quasi-Static Model

Conceptual Description   "Permanent" excess pore pressure is a decisive parameter for plastic deformation of normally consolidated cohesive soils under cyclic loading, see [1]. Both types of plastic deformation, undrained plastic shear deformation in the first stage and drained plastic volume strain in the second stage are dependent on this parameter. Upon this, a quasi-static model is proposed for describing the plastic deformation of saturated cohesive soils (normally to lightly overconsolidated) under cyclic loading, see Fig. 1. The individual cyclic loop can, however, not simulated by using this model.

As illustrated in Fig. 1, the maximum of cyclic stress σ_c is applied as a loading parameter. The Youngs modulus E is introduced to simulate the maximum reversible strain under cyclic loading. In principle, this part of strain changes over the course of each load cycle. However, it is normally much smaller than the permanent strain. Therefore E is assumed to be cycle number independent.

The sliding element in Fig. 1 represents shear and tension limits of saturated cohesive soils under cyclic loading. Test results from literature show that the dynamic frequency has influence on the shear strength. Under cyclic loading, the static shear strength of cohesive soils can be exceeded. However, this occurs mostly in the beginning of cyclic loading. The investigation of [2] using effective stress strength parameters indicated, that for normal consolidated clays, shear strength under undrained condition is not significantly affected by cyclic frequency and loading duration. Therefore it is assumed that static strength parameters can be applied for the cyclic loading case.

In the beginning of cyclic loading, most soil elements remain undrained and excess pore pressure rises depending on cycle number. The reversible part results from elastic deformation of soils and can be approximately assessed by using $u^c = A_c(\sigma_{x,c} + \sigma_{y,c} + \sigma_{z,c})$ (with $\sigma_{i,c}$ cyclic stress components). Parameter A_c is determined from cyclic undrained test. In comparison with u^c, the "permanent" excess pore water pressure u^b is caused by plastic shear deformation of the soil skeleton and is time-dependent. Using excess pore water pressure and total stresses (= static + quasi-static), failure strength upon the effective stress principle can be determined and plastic deformation ε^b occurs if the strength limit is exceeded.

Under cyclic loading, a particular aspect is the occurring plastic deformation of soils in the case that cyclic loading is inside the failure envelope. Here, plastic deformation consists of two parts. ε_u^b results from the undrained and ε_d^b from the drained part, see Fig. 1. As mentioned above, both of them are dependent on the "permanent" excess pore water pressure. The "permanent" excess pore pressure generated under undrained condition is treated as a pressure source $\psi = \partial u_g^b / \partial t$ and added to the consolidation equation,

$$\frac{\partial}{\partial x}\left(\frac{k}{\gamma_w}\frac{\partial u^b}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{k}{\gamma_w}\frac{\partial u^b}{\partial y}\right) + \frac{\partial}{\partial z}\left(\frac{k}{\gamma_w}\frac{\partial u^b}{\partial z}\right) = m_v\left(\frac{\partial u^b}{\partial t} - \psi\right) \quad (1)$$

The consolidation equation provides increasing excess pore pressure in the beginning of cyclic loading ($\Delta u^b > 0$), which is used for evaluating undrained plastic shear deformation through an empiric relationship. With the compressive volume coefficient m_v, the subsequent dissipation ($\Delta u^b < 0$) leads to drained plastic volume strain,

$$\Delta \varepsilon_d^b = m_v \cdot \Delta u^b \tag{2}$$

Excess pore water pressure under undrained cyclic loading condition From literature, e.g. [3] proposes that excess pore water pressure u_g^b under cyclic undrained condition can be evaluated by using,

$$\frac{u_g^b}{\sigma_{3,0}} = \rho \cdot \left(\frac{q_c}{\sigma_{3,0}}\right)^{\alpha_G} \cdot (1 + \log N)^\theta \tag{3}$$
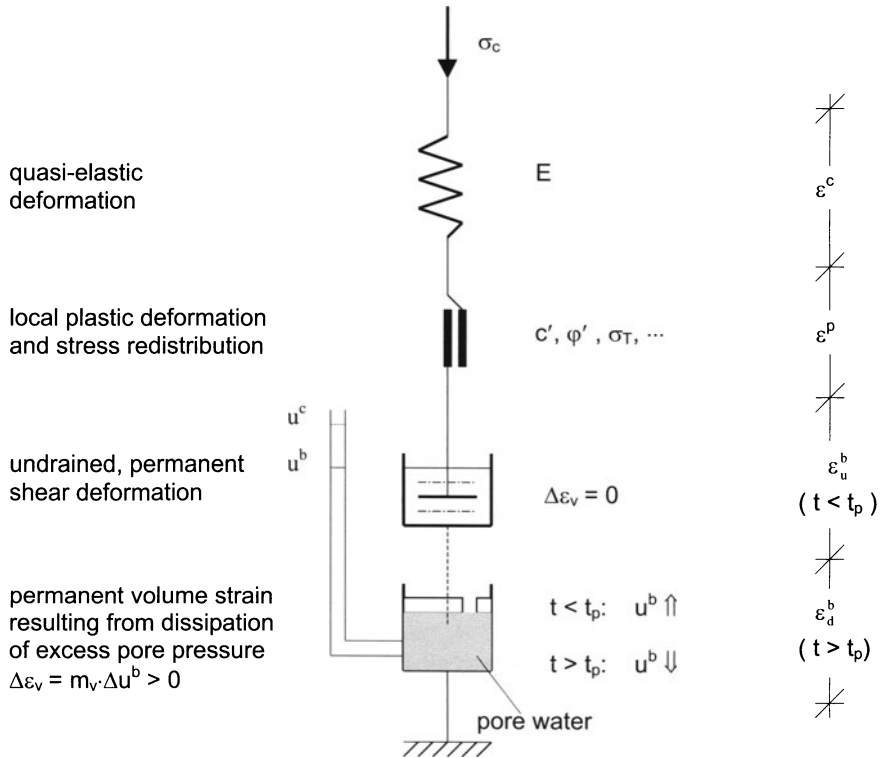
Fig. 1. Quasi-Static Model describing the Plastic Deformation of Saturated Cohesive Soils under Cyclic Loading

where ρ, α_G and θ are regression parameters from undrained cyclic tests. Test results of [4] with different frequencies (f = 0.02/0.05/0.2/0.5 Hz) indicate that by using the cycle number as a describing parameter in the analysis, higher excess pore water pressure is expected under low frequency condition. The difference probably arises from different loading time, as "permanent" excess pore water pressure is mainly caused by plastic shear deformation of the soil skeleton and therefore closely depends on the total loading time. In the case that test results are analysed using total loading time instead of cycle number, the influence of frequency could be seen as secondary. Upon this, the empiric relationship is rewritten as:

$$\frac{u_g^b}{\sigma_{3,0}} = \rho \cdot \left(\frac{q_c}{\sigma_{3,0}} \right)^{\alpha_G} \cdot (1 + \log t)^{\theta} \qquad (4)$$

Undrained Plastic Shear Deformation In strain theory, it is assumed that under undrained condition, shear sliding as well as partly collapse of soil skeleton lead to redistribution of loads between pore water and soil skeleton,

see [5]. The load originally carried by contact points of soil grains is partly transferred to pore water. As this shear deformation is predominately plastic, resulting excess pore water pressure is mostly "permanent". Hence a close relationship between undrained plastic shear strain and "permanent" excess pore water pressure can be assumed. For normally consolidated cohesive soils, there is an hyperbolic relationship between the ratio of excess pore pressure to consolidation pressure and the first principal strain, see [2]. It is independent of cyclic loading form and frequency and with the experimental parameters a and b written as,

$$\frac{u}{p_0'} = \frac{\varepsilon_{1,u}}{a + b \cdot \varepsilon_{1,u}} \tag{5}$$

Though the relationship given in Eq. 5 includes the reversible part of excess pore water pressure and strain, it could be concluded from strain theory that there is a similar relationship between irreversible parts of excess pore water pressure and strain. Therefore the following relationship is assumed,

$$\frac{u^b}{p_0'} = \frac{\varepsilon_{1,u}^b}{a + b \cdot \varepsilon_{1,u}^b} \tag{6}$$

Drained Plastic Volume Strain The parameter m_v in Eq. 1 and Eq. 2 represents the volume compressive coefficient under cyclic loading condition and should not be the same as that under static loading condition. It is a common way, however, to get this parameter by using undrained cyclic triaxial tests with subsequent dissipation. In this way, an empiric relationship between "permanent" excess pore water pressure u^b after undrained cyclic loading and plastic volume strain from dissipation can be established. The tests on Drammen clay (OCR=1) by [6] showed that plastic volume strain resulting from cyclic loading could be assessed by

$$\varepsilon_v^b = 1,5 \frac{C_r}{1 + e_0} \log\left(\frac{1}{1 - u^b/p_0'}\right) \tag{7}$$

C_r is the compressive parameter of conventional static oedometer tests in reloading stage. Upon this, the parameter m_v can be calculated using

$$m_v = d\varepsilon_v^b/du^b = 1,5 \cdot \frac{C_r}{1 + e_0} \cdot \frac{1}{\ln 10} \cdot \frac{1}{p_0' - u^b} \tag{8}$$

Using the relationships described above, the proposed quasi-static model in Fig. 1 has been numerically implemented (Program-System GeoCycl).

2.3   Verification of Numerical Model

For verification of the numerical model, the cyclic triaxial tests of [3] on Ariake clay (see Fig. 2) were reanalysed by using GeoCycl. Upon [3], samples were

fully consolidated under a vertical pressure of 59 kN/m². After installation of the soil samples in triaxial cells, they were statically loaded by a confining pressure of 100, 200 and 300 kN/m² respectively, and consolidated for 24 hours. Then, the samples were cyclicly loaded by a harmonic vertical load. In the course of cyclic loading, drainage was allowed laterally. Excess pore water pressure was measured in the center of the samples bottom. Based on the data published by [3], the parameters necessary for the numerical modeling are determined and illustrated in Fig. 2. The calculated results are compared with the test results of [3]. From this, it can be concluded that the developed numerical model predicts the development of excess pore pressure in the course of cyclic loading properly.
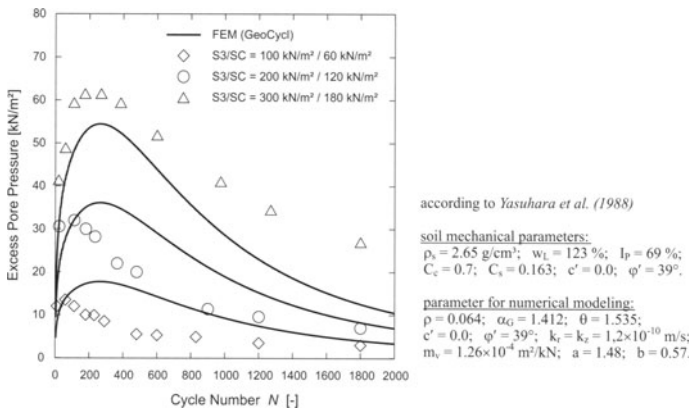
Fig. 2. Excess Pore Water Pressure: GeoCycl vs. Test Results of [3]

2.4   Cyclic Triaxial Tests on a Saturated Cohesive Soil

For the experimental investigation of plastic strain of cohesive soils under cyclic loading, cyclic triaxial tests on a clay of medium plasticity were carried out at Kassel University, see [1]. Soil samples of 10cm in diameter and 10cm in height were prepared by using Proctor compaction. They have a dry unit weight of 17.2 kN/m³ and a corresponding saturation grade of 95%. The soil mechanical indexes are given in Fig. 3. After installation of the prepared soil sample into the triaxial cell and the cellfilling with demineralised water, the saturation procedure (5 days) with a back pressure of 300 kPa was applied. Then, the soil samples were consolidated. After the completion of consolidation, the soil samples were vertically loaded by cyclic dynamic stress. In the course of cyclic loading, drainage was allowed on the circular surface of the samples. The loading conditions and procedure are illustrated in Fig. 3. The analysis of the test results showed that plastic strain reached a stable

soil mechanical parameters

$w_L = 41.9\%$ $w_P = 18.4\%$
$w_n = 22.7\%$ $I_P = 23.5\%$
$I_C = 0.82$ $Gs = 2.75$
$w_{pr} = 16.4\%$ $\rho_{pr} = 18.1 kN/m^3$
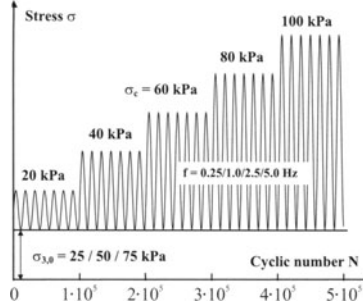
Fig. 3. Loading Conditions and Procedure of the Cyclic Triaxial Tests on Clay of Medium Plasticity

state after about 105 cycles in each load step. In Fig. 4, plastic strain ratio is illustrated depending on the cycle number (a) and on the loading duration (b) for the case of $\sigma_{3,0} = 50 kN/m^2$, $\sigma_c = 60 kN/m^2$ and $f = 0.25 Hz$ / $2.5 Hz$. A deviation between the plastic strains at different frequencies is evident by using cycle number N, while both curves become almost the same, when loading duration t is used as a describing variable. This indicates that loading duration t may be controlling parameter for determining the plastic strain under cyclic condition. In Fig. 5a, the plastic axial strain in stabile state is
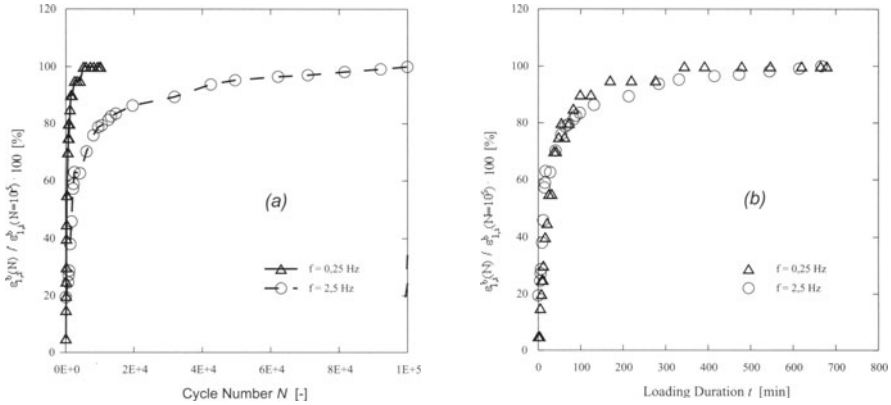
Fig. 4. Plastic Axial Strain as a Function of (a) Cycle Number N and (b) Loading Duration t for the Case $\sigma_{3,0} = 50 \ kN/m^2$, $\sigma_c = 60 \ kN/m^2$ and $f = 0.25 Hz$ / $2.5 Hz$

expressed as a function of $\sigma_c/\sigma_{3,0}$ for the cases $\sigma_{3,0} = 25/50/75 \ kN/m^2$ and $f = 1$ Hz. Test results at different static and dynamic stresses nearly provide the same curve, approximated by,

$$\varepsilon_p^a = \alpha \cdot \left(\frac{\sigma_c}{\sigma_{3,0}}\right) + \beta \cdot \left(\frac{\sigma_c}{\sigma_{3,0}}\right)^2 \tag{9}$$

with α and β: as the regression curve coefficients. For the case of $\sigma_{3,0} = 50$
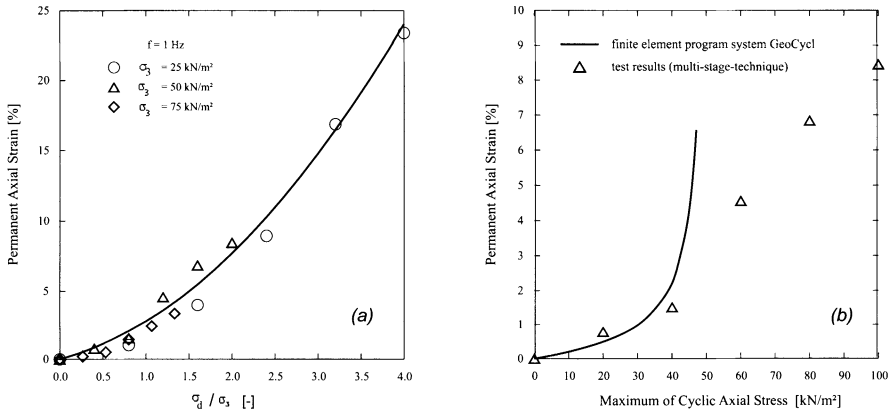
Fig. 5. (a) Plastic Axial Strain in Stabile State as a Function of $\sigma_c/\sigma_{3,0}$ and (b) Comparison of Numerical and Experimental Results for the Case $\sigma_{3,0} = 50kN/m, f = 1Hz$

kN/m^2 and $f = 1$ Hz, numerical calculations by using GeoCycl were carried out. Results are shown in Fig. 5b. In the first two steps, good agreement is achieved. Large deviation occurs with the third step, which indicates a stiffer experimental behavior of the soil sample against the theoretical predictions. This may be put down to the strengthening effect resulting from the multi-stage technique, see Fig. 3. However, this effect is not included in the quasi-static model.

3   Granular Soils and Ballast

3.1   Experimental Investigation and Derivation of a Cyclic-Viscoplastic Constitutive Approach

General Based on the results and knowledge gained from the European BRITE/EURAM III research project EUROBALT II, further research within the framework of the DFG-priority program has been carried out regarding the plastic and elastic behavior of granular soils and ballast under cyclic-dynamic loading. Aiming at the constitutive description of the elastic and plastic long-term behaviour, the setup of functional correlations, in dependence of the essential parameters of influence is intended.

$$E_{S,dyn} = \frac{\sigma_{dyn}}{\varepsilon_{ce}} = f(\sigma_3, \sigma_{dyn}, N, f_B, e, \varphi') \qquad (10)$$

$$\varepsilon_{cp} = f(\sigma_3, \sigma_{dyn}, N, f_B, e, \varphi') \qquad (11)$$

Whereas $E_{S,dyn}$ is the dynamic secant modulus, ε_{cp} the cyclic-plastic strain, σ_{dyn} the applied dynamic stress, σ_3 the confining stress, f_B the loading frequency, N the load-cycle number, e the void ratio and φ' the static friction angle. Departing from these investigations, a constitutive approach for granular soils and ballast is derived, which basically is an advance from the quasi-static modell of [1].

Testrig and Investigated Material From Fig. 6 a schematic cross section of the newly developed combined *C*yclic *T*riaxial *R*esonant *C*olumn test device (CTRC) can be seen. The CTRC test rig was developed within the framework of the EUROBALT II research project. The major advantage of the CTRC is the ability of investigating simultaneously both, the plastic long-term behavior with increasing load-cycle numbers, as well as the change of the elastic behavior with the ongoing loading history. Another advantage of the CTRC is to be seen in the dimensions of the installed sample, with 500mm in diameter and a height of 1000mm. With such dimensioned sample the investigation of real scale railway ballast 22.4/63 is made possible. An in depth description of the CTRC test rig is to be found in [7]. From Fig. 7
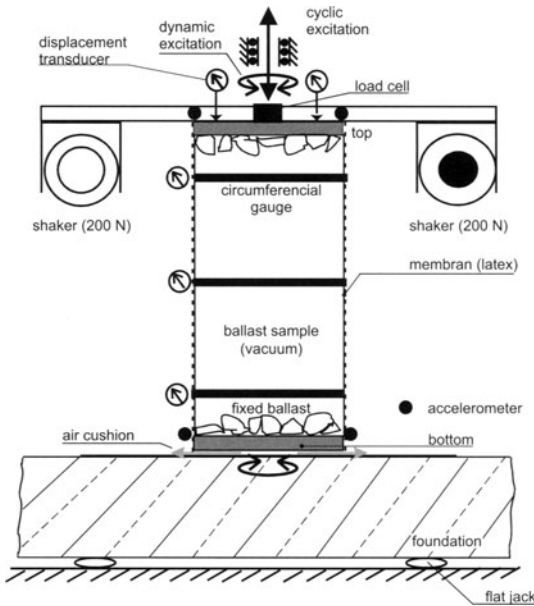
Fig. 6. Cross Section of the *C*ombined *C*yclic *T*riaxial *R*esonant *C*olumn Test device (CTRC)

the grain-distribution curves of investigated soils and ballast materials in the CTRC can be seen.
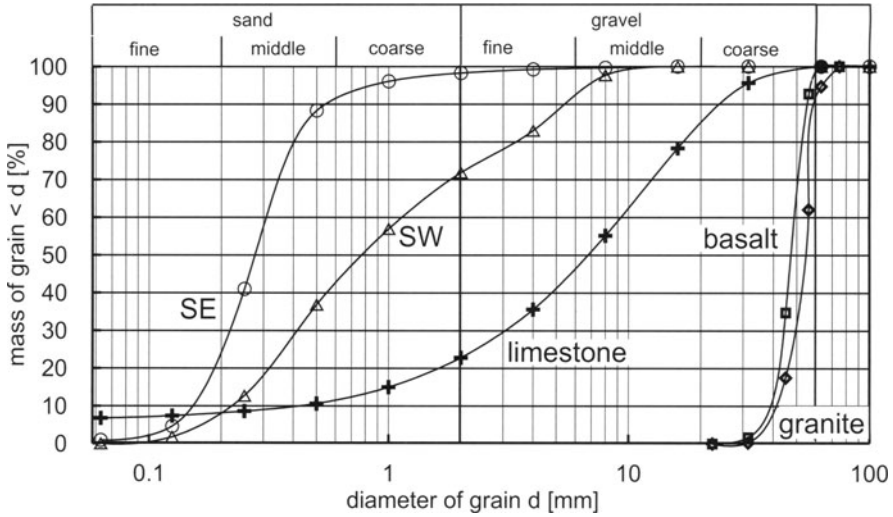
Fig. 7. Grain-Distribution of Investigated Soils and Ballast Materials

Plastic Long-Term Behaviour Results The cyclic loading of the cylindrical sample has been realised with a harmonious sine function with N = 1E+05 to 1E+06 load cycles. The data aquisition takes place in time windows with a high sample frequency at distinct time intervals. With such procedure, the elastic and plastic strains can be identified and distinguished. Hence it is possible to determine, next to the plastic strains, the changing elastic behavior vs. load-cycle number. An in depth description of the data acquisition and evaluation of CTRC test results is to be found in [8]. In the following, the general elastic and plastic material behaviour under cyclic-dynamic loading is outlined. From Fig. 8a the plastic strain ε_{cp} can be seen in a half-logarithmic representation. There is a nonlinear behaviour to be observed, with a dependency onto the cyclic dynamic loading. The standardised cyclic deviatoric stress ratio X is given with,

$$X = \frac{(\sigma_1 - \sigma_3)_c}{(\sigma_1 - \sigma_3)_{s,f}} \tag{12}$$

Whereas $(\sigma_1 - \sigma_3)_c$ is the cyclic-dynamic deviatoric stress and $(\sigma_1 - \sigma_3)_{s,f}$ the static deviatoric failure stress in triaxial condition. From Fig. 8b the elastic behaviour of the investigated granular soil, in terms of the dynamic secant modulus, can be seen. Fig. 8b depicts the increase of the dynamic secant modulus with progressing cycle-numbers in a double-logarithmic representation. Partly the increase amounts up to 100% after 1E+06 load-cycles. The increase of stiffness comes along with the decrease in elastic strain and results from the contractant behaviour of the granular soils as well as the bracing of the grain structure.
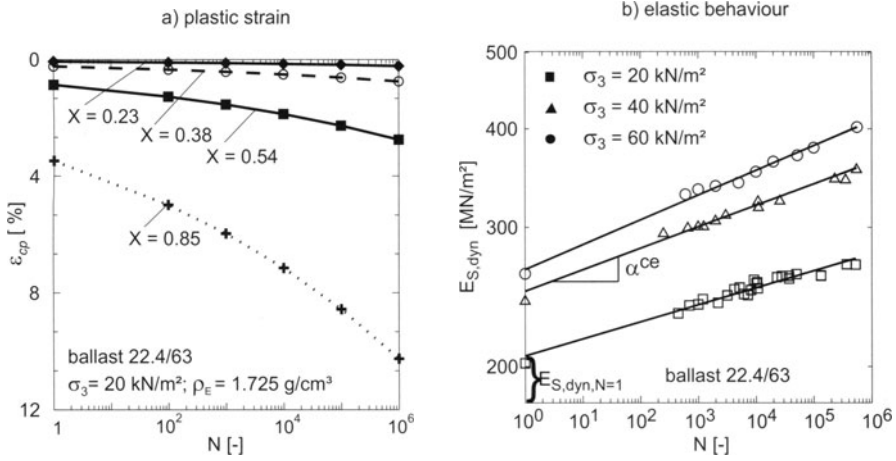
Fig. 8. Plastic and Elastic Behaviour under Cyclic-Dynamic Loading

From Fig. 9a the evolution of volumetric strain vs. cycle number can be seen. Up to a certain cycle number, N_{crit}, independent of compaction during installation, all test samples respond to the cyclic excitation with dilatant behaviour. Afterwards the sample experiences compression (contractancy), see Fig. 9a. The transition from dilatant to contractant behaviour for the investigated samples, lies inbetween cycle numbers $N_{crit} = 1$ to $N_{crit} = 100$. From the semi-logarithmic representation of $\Delta V/V$, the mean value could be determined as $N_{crit} = 10$. The evolution of volumetric strain and void ratio have to be seen in close conjunction, with a qualitatively similar behaviour.

From Fig. 9b the evolution of void ratio e vs. cycle number N can be seen. In this representation, a maximum of the void ratio e can be identified at $N_{crit} = 10$. Up to N_{crit}, the void ratio e increases. Having reached N_{crit}, the pore number decreases continuously, which is a contractant behaviour. The maximum attainable void ratio under cyclic-dynamic loading is called, similar to static traixial tests, a "pseudo critical void ratio $e_{ps,crit}$". It describes the maximum volumetric expansion in the cyclic-dynamic triaxial test.

Derivation of a new Cyclic-Viscoplastic Constitutive Approach The plastic and elastic stress-strain cycle-number behaviour can be described by means of the modified viscoplastic model in Fig. 10. The total cyclic-dynamic strain $\varepsilon_{c,tot}$ is composed by the single plastic and elastic portions,

$$\varepsilon_{c,tot} = \varepsilon_{cp} + \varepsilon_{ce} \tag{13}$$

The extended viscoplastic model represents a modification of the classic elastic-viscoplastic model and the quasi-static theory of [1]. Hence it is called a *cyclic-viscoplastic constitutive approach* (CVP-model). The cyclic-visco-
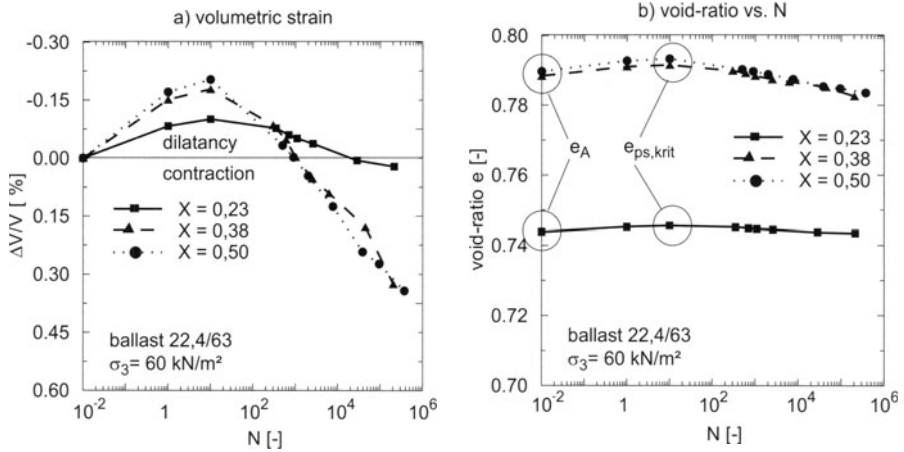
Fig. 9. Volumetric Strain and Void Ratio vs. Cycle Number of Cyclicly-Dynamically loaded Samples
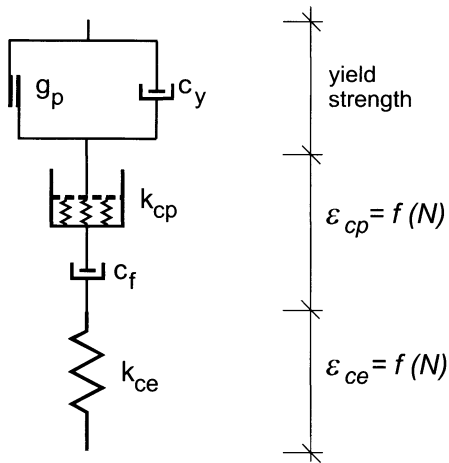
Fig. 10. Simplfied Viscoplastic and Cyclic-Viscoplastic Rheological Model

plastic model theory is composed of two constitutive approaches, the cyclic-viscoplastic and the cyclic-viscoelastic constitutive approach. The separate model elements are explained in depth in [8].

In the following the derivation of the cyclic-viscoplastic constitutive approach is given. By means of a power functional, the plastic strain is written as,

$$\varepsilon_{cp,N} = \varepsilon_{cp,N=1} \cdot N^{\alpha_{cp}} \tag{14}$$

Whereas $\varepsilon_{cp,N=1}$ is the plastic strain for the first load cycle and α_{cp} the linear increase for progressing cycle numbers in a double-logarithmic representation.

As a result of the accomplished cyclic-dynamic triaxial tests, it has been possible to set the equation-parameters of Eq. 14 against those of Eq. 10 and Eq. 11 and derive corresponding equation systems.

On the basis of this analysis, $\varepsilon_{cp,N=1}$ can be written as,

$$\varepsilon_{cp,N=1} = \phi_f \cdot \kappa_p \cdot e^{\lambda_p} \tag{15}$$

whereas ϕ_f is a functional of the loading frequency, e the void ratio and κ_p as well as λ_p another functions, mapping the influence of the confining stress σ_3, the dynamic stress σ_{dyn} and the shear strength φ'. The void ratio e of granular materials has been identified as the most essential material parameter for the description of the cyclic-plastic strain $\varepsilon_{cp,N=1}$ at first load cycle.

In a similar way, a functional for the description of the increase α_{cp} is given,

$$\alpha_{cp} = \kappa_p^\alpha \cdot \exp^{(-\lambda_p^\alpha \cdot f_B)} \tag{16}$$

The increase α_{cp} is dependent on the state of stress, the stress amplitude as well as the loading frequency. The influence of the loading frequency can be seen from Eq. 16. With increasing loading frequency f_B, there is a decrease in the cyclic-plastic strain increments. The curve increase α_{cp} is less steep with higher loading frequency. The porosity of the granular material on the other hand has no influence on the increase α_{cp} of the cyclic-plastic strain.

For a detailed description of the derivative of the cyclic-viscoplastic constitutive apprach with all it's plastic and elastic portions, see [8]. In the following focus is set rather on the numerical implementation into the commercial FE-code ABAQUS, as well as a presentation of first numerical modelling results.

3.2   Numerical Implementation

General Currently, the numerical modelling of geotechnical materials under cyclic-dynamic loading can be differentiated as follows:

- dynamic modelling of the linear elastic response of the investigated structure due to the induced excitation, "*short-term dynamics*", and
- modelling of the accumulating remaining strains over the long-term, "*long-term (deformation) behaviour*"

Modelling activities carried out in this work are focused on the long-term / settlement behaviour, the continuous explicit time/load-cycle dependent degradation of foundations and track systems, caused by deformations from all corresponding layers below.

Theoretical Approach To exemplify, from Fig. 11 the evolution of strain vs. time/load-cycles can be seen. Generally the solution techniques are differentiated into implicit and explicit formulations. In this context, the terms implicit and explicit do not have to be mistaken with the well known concept of strain rate integration in the theory of classic plasticity. The implicit
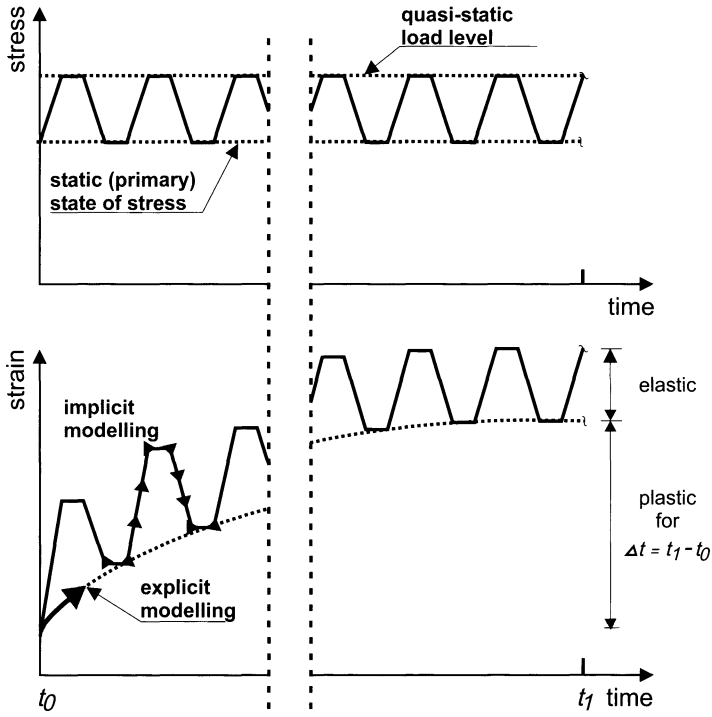
Fig. 11. Solution techniques, implicit vs. explicit formulation

technique describes the whole structure with the dynamic phenomena (in the short-term) and the long-term deformation behaviour in one global model. Thus the dynamic behaviour in an incrementally changed structure has to be simulated continuously/implicitly over the course of each single cycle, see Fig. 11, until a relevant remaining deformation has occurred. However such deterioration or deformation normally result from higher load-cycle numbers only. Also the constitutive material behaviour itself must be treated in a nonlinear manner. Obviously the dynamic modelling of the hysteretic material behaviour is quite complicated from a theoretical point of view and for the required high number of load-cycles also inconvenient from a practical point of view. On the one hand adequate theoretical models are missing or are just too complicated for practical applications. On the other hand today's computing capacities do not seem to be sufficient for such extensive problems.

Within this work a macroscopic explicit formulation technique has been chosen, where single amplitudes of some quantity at discrete times, or the course of a single cycle in real time do not play the determining role. Certainly, focus is set on the repeated loading with it's resulting accumulation of permanent strain/deformation for a finite number of cycles. Such approach is of practical meaning for engineers, in spite of the theoretically required exact course of each cycle, see Fig. 11. Continuous modelling of granular soil otherwise becomes very difficult if not at all impossible. Therefore the loading scheme is applied in a static manner and due to the reflection of the cyclic dynamic influences onto the material deformation behaviour within the constitutive routine, the modelling is described as quasi-static. Such procedure simplifies the constitutive description, as forces of inertia and the whole problems connected with the dynamic wave propagation in media don't have to be taken into consideration. In general assumptions and simplifications have to be done to eliminate the impossibility of giving respect to all occurring effects in reality. Hence, the effects of pore water pressure and liquefaction processes are not included in the approach.

Cyclic-Viscoplastic Strain Increments To be enabled to implement the constitutive approach described in Sec. 3.1, criteria for the stabilisation versus the progressive failure of soil under cyclic loading have to be defined. For the numerical model an analogous stress-strain relationship, observed in cyclic triaxial tests, is taken into consideration. Depending on the consolidation pressure, there are two deformation modes observed, see Fig. 12:

- Cyclic Calming and
- Cyclic Failure.

For the distinction / separation of the two states of material behaviour, a "reduced", i.e. modified cyclic Mohr Coulomb failure criterion F_c is applied. For
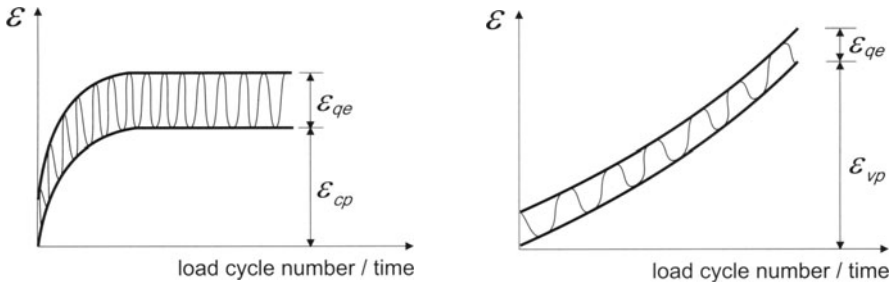
Fig. 12. Typical deformation modes observed in cyclic triaxial tests

cyclic states of stress satisfying the failure criterion F_c, the resulting cyclic permanent strains ε_{cp} of the soil specimen will converge vs. a stable value.

In this state of cyclic stress the soil specimens behaviour can be described as quasi-elastic. Consequently this deformation mode is called "Cyclic Calming". For cyclic states of stress violating the criterion, the soil specimen in the triaxial test will show signs of progressive failure. Such behaviour, in the model realised analogous to the theory of viscoplasticity, hereafter is called "Cyclic Failure". A local overshoot of the failure criterion, signified by $F_c > 0$ at cycle number N, will cause, next to the cyclic-plastic strain, additional viscoplastic straining, as long and at a rate that is related to the amount by which the yield has been violated. Such a quasi-static model for granular materials and ballast under cyclic loading has been derived, and by means of a cyclic-viscoplastic constitutive approach implemented into the FEM program ABAQUS. In Fig. 13 the constitutive approach, as it has been implemented into ABAQUS, has been depicted in the principal stress space. According to
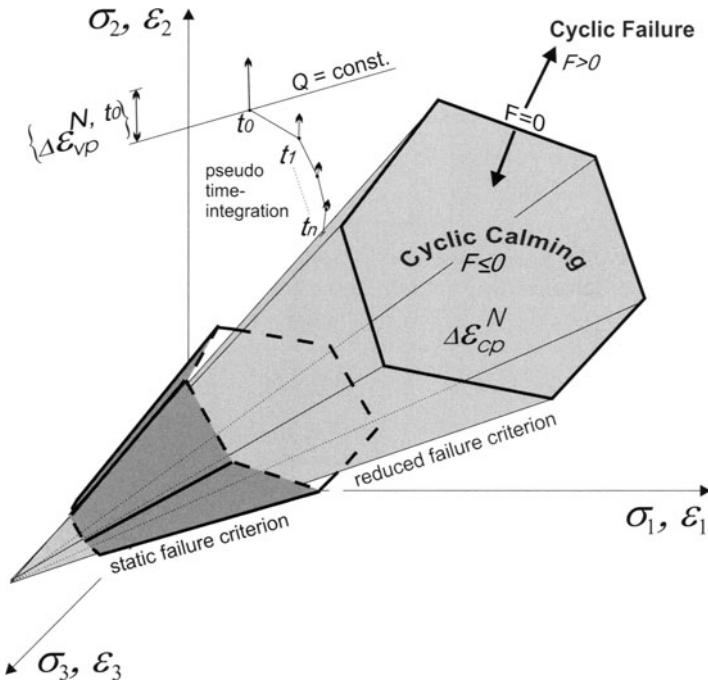
Fig. 13. Rheological model for granular materials and ballast under cyclic loading for the principal stress space

the constitutive approach, the incremental evolution equations for the cyclic-plastic strains are formulated for the principal stress space. Hence, principal strain increments, the equivalent to those occurring in a cyclic triaxial test, have to be rotated into the local axis system at integration point level.

Main assumptions for the implementation of the constitutive approach are:

- the cyclic loading can be idealised in a quasi-static manner with a static stress level
- there are effective stresses only / no liquefaction phenomena (no pore water pressures)
- the "Cyclic Failure" can be described by means of a classical viscoplastic formulation / initial strain method
- the incremental evolution of the "Cyclic Calming" can be described by means of a variation of double logarithmic functions derived in cyclic triaxial tests

3.3   Modelling Examples

Verification Example Exemplary, a numerical simulation of cyclic triaxial tests on ballast material is documented in the following. From Fig. 14 a comparison of the cyclic-plastic strains, measured in the CTRC test rig, and those calculated with ABAQUS can be seen. From the results plot, the
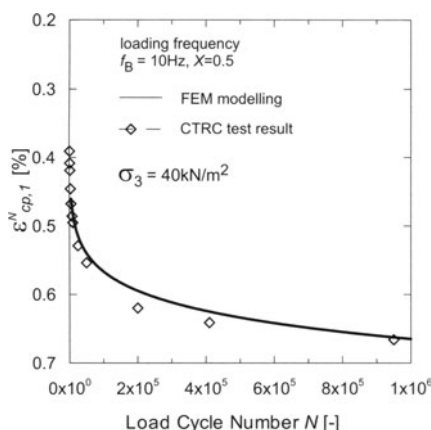
Fig. 14. Measured cyclic-plastic strains in the CTRC test rig vs. numerical modelling results with ABAQUS

comparison of the axial cyclic-plastic strain for a primary state of stress $\sigma_3 = 40\text{kN/m}^2$ can be seen. There is a good correlation (coefficient of correlation r $= 0.985$) between the modelling results with ABAQUS and those obtained in the CTRC test in laboratory. Generally the numerical modelling of principal cyclic-plastic strains of the Cyclic Calming have proven a good congruency to the CTRC test results, representing the axial and radial strains.

Bridge End Transition Zone The following modelling example is based on a real in situ bridge end transition zone. The bridge end construction is designed with a cement reinforced soil wedge placed behind the bridge, see Fig. 15. The backfill and frost protection layer material consist of gravely material with a very high compaction. From suitable soil samples taken on site, cyclic-viscoplastic material parameters have been determined in laboratory by means of geotechnical standard tests as well as sequences of cyclic triaxial testing. Furthermore, by means of comparative numerical modelling of large scale dynamic load plate tests, carried out on site, the cyclic viscoplastic material parameters have been verified for modelling of the transition zone. At
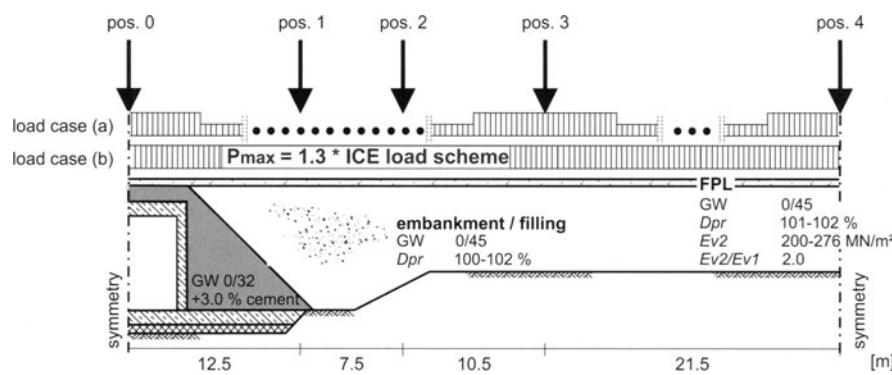
Fig. 15. Bridge end transition zone, geometry, materials and load cases

the time of modelling, the in situ track had been completed up to the top of the frost protection layer. As no information had been given, a slab track superstructure type Rheda had been chosen for the modelling, see Fig. 16. Furthermore from the cross section it can be seen, that the discretised geometry in the FE model has been modified. The symmetry axis has been moved to the left to guarantee the 3D load distribution and at the same time keep the model to a straightforward size. Due to the quasi-static nature of the numerical modelling, there are 2 load cases to be differentiated, see Fig. 15. Load case (a) is determined by a UIC 71 equivalent load scheme, with the exception of the load level Pmax derived from the actual ICE train axle loads and geometry. To take the dynamic load amplification into consideration, the value of P is charged with a dynamic factor of k = 1.3. Within the variation of load case (a), the load scheme itself is moved alongside the track axis (pos. 0 to 4), hence allowing the investigation of evolving settlement dells. Considering that a stationary quasi-static load scheme can not be the true image for a moving train load, the infinite load case (b) with Pmax is taken into consideration. Load case (b) is seen as the worst case scenario and as an upper bound for the settlements alongside the track. From Fig. 17 the numerical modelling results for the different load cases and load scheme po-
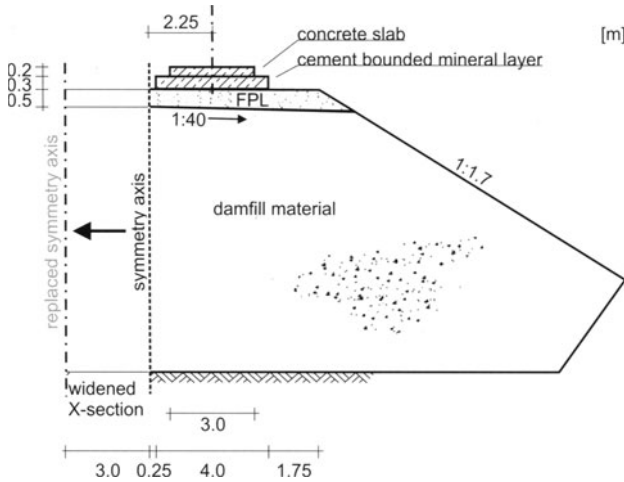
Fig. 16. Cross section bridge end transition zone, slab track superstructure type Rheda

sitions can be seen for a load-cycle number of $N = 4.9E+07$. Such number is an equivalent for a mean track loading in 60 years. In accordance with the given expectation, the infinite load case (b) represents an upper bound for the cyclic-plastic settlements. Although the absolute values of the settlement dells, load case (a), come close to the settlement along the track, load case (b), further investigations can be limited to the latter one. In any case there will no such dells be observed in situ.
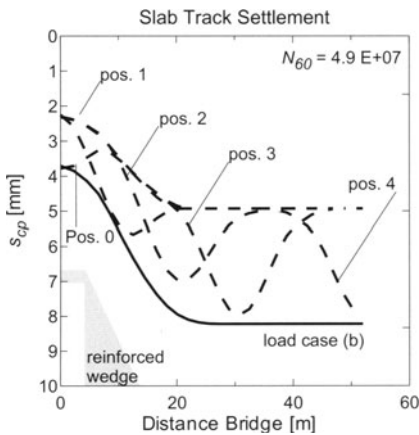
Fig. 17. Modelling results, bridge end transition zone, settlements for load cases (a) and (b) after load-cycle number $N = 4.9E+07$.

4   Outlook and Future Objectives

The executed investigations on water saturated soft soils under cyclic loading have to be further validated experimentally. In particular the evolution of pore water pressure under cyclic loading is to be verified and the theoretical approaches to be improved.

The presented cyclic-viscoplastic constitutive approach for granular and ballast materials has proven to be an appropriate enhancement for the quasi-static modelling of such materials under cyclic-dynamic loading. The objective for future research work is seen in further improvements of the constitutive approach, especially the enhancement to further gradings of granular materials. Another objective is the modelling itself, i.e. parametric studies of complex in situ situations, allowing prediction of deterioration behaviour as well as general improvements of construction guidelines. Finally the derivation of a corresponding quasi-static approach for joint behaviour in numerical interface modelling has to be undertaken.

References

1. Hu, Y. (2000) Zum Verformungsverhalten von wassergesättigten bindigen Böden unter zyklischer Belastung. Schriftenreihe Geotechnik, Heft 8, Universität Kassel
2. Yasuhara, K., Yamanouchi, T. et al. (1982) Cyclic Strength and Deformation of Normally Consolidated Clay. Soils and Foundations, Japanese Society of Soil Mechanics and Foundation Engineering, 22(3)
3. Yasuhara, K., Hirao, K. et al. (1988) Partial-Drained Behaviour of Clay under Cyclic Loading. In: Proc. 6th Int. Conf. on Numerical Methods in Geomechanics, Innsbruck, Swoboda (ed.)
4. Matsui, T., Ohara, H. et al. (1980) Cyclic Stress-Strain History and Shear Characteristics of Clays. J. Geotech. Engrg., ASCE, 106 (10)
5. Lo, K. Y. (1969) The Pore Pressure - Strain Relationship for Normally Consolidated Undisturbed Clays, Part I-II Theoretical Considerations. CGJ (6)
6. Yasuhara, K. and Andersen, K. H. (1991) Recompression of Normally Consolidated Clay after Cyclic Loading. Soils and Foundations, 31(1)
7. Gotschol, A., Kempfert, H. G., Meyer, R., Stöcker, T. (2000) Bodendynamisches Versuchsgerät zur kombinierten Bestimmung zyklischer und dynamischer Kenngrössen granularer Böden und Schotter, Geotechnik 23, Heft 1, S. 53-58
8. Gotschol, A. (2002): Veränderlich elastisches und plastisches Verhalten nichtbindiger Boeden und Schotter unter zyklisch-dynamischer Beanspruchung. Schriftenreihe Geotechnik, Heft 12, Universität Kassel
9. Stöcker, T. (2002): Zur Modellierung von granularen Materialien bei nichtruhenden Lasteinwirkungen. Schriftenreihe Geotechnik, Heft 13, Universität Kassel

3D-Simulation of Dynamic Interaction Between Track and Layered Subground

Stavros A. Savidis, Reinhold Hirschauer, Christopher Bode, and
Winfried Schepers

Technische Universität Berlin, Fachgebiet Grundbau und Bodenmechanik,
Sekr. TIB1-B7, Gustav-Meyer-Allee 25, D-13355 Berlin, Germany

Abstract. Three-dimensional numerical models to analyze dynamic soil–structure
interaction (SSI) of rail tracks on layered soil have been developed. For track struc-
tures on the surface we implemented a frequency domain as well as a time domain
approach. To incorporate soil irregularities we also implemented a frequency do-
main approach for embedded structures. A common feature of all models is the
application of the substructure method. The track or any other structure is ana-
lyzed be means of finite elements, whereas the unbounded soil is analyzed by means
of boundary elements with layered halfspace fundamental solutions. To provide the
user with all the capabilities of a powerful commercial finite element software pack-
age the SSI-models have been implemented in ANSYS. Therefore nonlinear prop-
erties of the track and of the soil up to some extent can be taken into account using
standard ANSYS features.

1   Introduction

The aim of the project was the development of three-dimensional numerical
models to analyze soil–structure interaction of rail tracks on layered soil under
dynamic load. The following configurations were considered:

- structure on surface analyzed in frequency domain,
- structure on surface analyzed in time domain,
- embedded structure analyzed in frequency domain.

A common feature of all models is the application of the substructure method.
The track or any other structure is analyzed be means of finite elements,
whereas the soil is analyzed by means of boundary elements to fully satisfy
the radiation condition towards infinity.
The models have been implemented in ANSYS to provide the user with all
the capabilities of a powerful commercial finite element software package.
Therefore nonlinear properties of the track and of the soil up to some extent
can be taken into account using standard ANSYS features, whereas FOR-
TRAN code for the analysis of the soil has been developed and linked to
ANSYS using its UPF – User Programmable Features – interface.
In the following sections we will first describe the substructure method at
a glance, followed by a more detailed description of the above mentioned

configurations. For each configuration we describe the analysis of the track, the specific application of the substructure method to analyze the soil, and the coupling, followed by some applications.

2   Substructure Method

By using the substructure method we devide a complex structure into several parts. We keep one part as a kind of "main structure" and cut off the substructures (Fig. 1). All substructures can be analyzed independently with respect to the degrees of freedom (DOF) at the common interface with the main structure. By satisfying equilibrium and compatibility conditions at all interfaces the whole structure is reassembled and a solution for the complete system can be obtained.
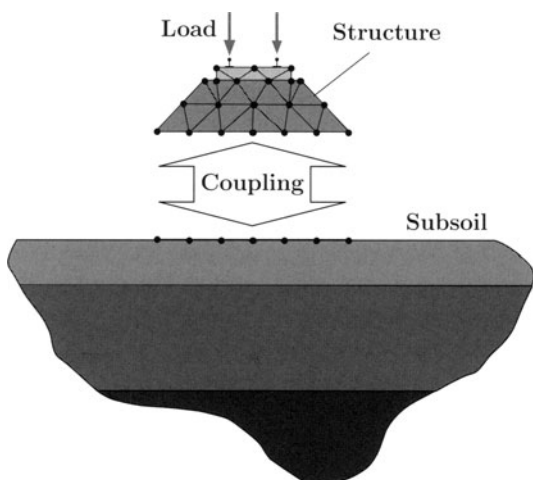
Fig. 1. Substructure method

Advantages of the substructure method:

- for different physical properties specific techniques can be applied for the analysis to retain the advantages of one method and to eliminate the disadvantages of the respective others,
- reduction of memory requirements since small systems are analyzed,

Disadvantage of the substructure method:

- to obtain the displacements and stresses at arbitrary points of a substructure, a post-solution step must be performed with the displacements and stresses at the interface of the substructure as input.

In particular the first advantage listed above was used in this work. Finite elements were applied for the track and partly also for the surrounding soil, which may exhibit nonlinear properties, while boundary elements with halfspace–fundamental solutions were used for the unbounded soil.

3   Structures on the Surface

In this section we give a more detailed description of the substructure method applied to structures on the surface of the layered soil in frequency domain as well as in time domain.

3.1   Analysis in Frequency Domain

The time dependence of all variables is assumed to be harmonic of type

$$A(t) = \hat{A} \cdot e^{i\omega t} \ . \tag{1}$$

Here $\hat{A}$ denotes a complex amplitude of any kind, t is the time, ω is the circular frequency and i is the imaginary unit. For simplicity the exponential term as well as the hat ($\hat{\ }$) is omitted in the following. Different time dependencies will be denoted explicitly.

Structure

For modelling a finite structure – in this case the track – the finite element method is used here since it is well established for a wide range of problems in engineering.

The discrete equation of motion by means of finite elements is given as

$$\boldsymbol{M} \cdot \ddot{\boldsymbol{u}}(t) + \boldsymbol{C} \cdot \dot{\boldsymbol{u}}(t) + \boldsymbol{K} \cdot \boldsymbol{u}(t) = \boldsymbol{P}(t) - \boldsymbol{Q}(t) \ . \tag{2}$$

The vector $\boldsymbol{u}$ is the displacement vector, $\dot{\boldsymbol{u}}$ the first and $\ddot{\boldsymbol{u}}$ the second derivative with respect to time. The matrices $\boldsymbol{M}$, $\boldsymbol{C}$, and $\boldsymbol{K}$ are the mass matrix, damping matrix, and stiffness matrix of the structure, respectively. The vector $\boldsymbol{P}$ denotes the external forces and $\boldsymbol{Q}$ denotes the resultant nodal forces of the contact stresses at the interface of track and soil. Assuming harmonic time dependency as in (1) we obtain

$$(-\omega^2 \boldsymbol{M} + i\omega \boldsymbol{C} + \boldsymbol{K}) \cdot \boldsymbol{u} = \boldsymbol{P} - \boldsymbol{Q} \tag{3a}$$

$$\text{or}$$

$$\boldsymbol{S} \cdot \boldsymbol{u} = \boldsymbol{P} - \boldsymbol{Q} \tag{3b}$$

with $\boldsymbol{S}$ the complex dynamic stiffness matrix of the structure.

The DOFs of (3b) are sorted by DOFs at the interface (subscript I) and the remaining DOFs outside the interface (subscript R) and we obtain

$$\begin{bmatrix} \boldsymbol{S}_{\mathrm{RR}} & \boldsymbol{S}_{\mathrm{RI}} \\ \boldsymbol{S}_{\mathrm{IR}} & \boldsymbol{S}_{\mathrm{II}} \end{bmatrix} \cdot \begin{Bmatrix} \boldsymbol{u}_{\mathrm{R}} \\ \boldsymbol{u}_{\mathrm{I}} \end{Bmatrix} = \begin{Bmatrix} \boldsymbol{P}_{\mathrm{R}} \\ \boldsymbol{P}_{\mathrm{I}} \end{Bmatrix} - \begin{Bmatrix} \boldsymbol{0} \\ \boldsymbol{Q}_{\mathrm{I}} \end{Bmatrix} \tag{4}$$

For simplicity we use the symbols $\boldsymbol{Q}$ and $\boldsymbol{Q}_{\mathrm{I}}$ interchangeable, because by definition the contact stresses only exist at the interface. The dynamic stiffness matrix $\boldsymbol{S}$ in (3b) and (4) is subject to standard finite element procedures and is therefore computed with ANSYS. The remaining task is the computation of $\boldsymbol{Q}$, which is described in the following section.

Subsoil

Application of boundary elements [10] requires fundamental solutions which satisfy the differential equation of the investigated system, but not necessarily the given boundary conditions.

Usually fundamental solutions for a homogeneous fullspace are used. Since they don't satisfy the layered halfspace boundary conditions, they require – strictly speaking – placement of an *infinite* number of boundary elements on the surface and even at all layer boundaries. Therefore one must carefully choose a distance from the interface where to truncate the discretization such that the influence of the violated boundary conditions of the fundamental solutions becomes negligible.

To avoid these obstacles we use fundamental solutions which satisfy exactly the traction free boundary condition at the surface of a halfspace and also consider horizontal layering. Thus, it's sufficient to discretize the soil solely at the soil–structure interface.

For homogeneous soil we use Lamb's solution [20], while for a layered halfspace we use the Thin Layer Method (TLM) [18,19], a powerful semianalytical solution, which includes material damping and all effects due to the layering like reflection and refraction at layer boundaries, dispersion, and geometric damping.

Standard boundary element schemes yield to an integral equation [10]

$$c_{ij}(\boldsymbol{\xi})\, w_j(\boldsymbol{\xi}) + \int_{\Gamma} q_{ij}^*(\boldsymbol{x},\boldsymbol{\xi})\, w_j(\boldsymbol{x})\, d\Gamma = \int_{\Gamma} w_{ij}^*(\boldsymbol{x},\boldsymbol{\xi})\, q_j(\boldsymbol{x})\, d\Gamma \;, \qquad (5)$$

where w_j and q_j denote the unknown displacements and stresses of the soil, w_{ij}^* and q_{ij}^* denote the fundamental solutions in terms of displacements and stresses in the j direction due to a unit point load acting at $\boldsymbol{\xi}$ in the i direction, $\boldsymbol{x}$ is an arbitrary point on the boundary Γ, and c_{ij} is the discontinuity jump term and depends on the smoothness of the boundary at point $\boldsymbol{\xi}$.

Since our fundamental solutions satisfy the traction free boundary conditions at the halfspace surface, the integral on the left hand side of (5) vanishes completely at all non–singular points. Furthermore, it can be proven that in our case the left hand side is equal to $\delta_{ij}\, w_j(\boldsymbol{\xi}) = w_i(\boldsymbol{\xi})$. Thus, after discretization with boundary elements and applying (5) to each nodal point $\boldsymbol{\xi}$ – the so-called interaction points – we obtain the linear equation system

$$\boldsymbol{w} = \boldsymbol{F} \cdot \boldsymbol{q} \qquad (6)$$

with $\boldsymbol{F}$ the flexibility matrix, $\boldsymbol{w}$ the displacements and $\boldsymbol{q}$ the contact stresses at the interaction points.

It's worth mentioning that in (6)

- all quantities are frequency dependent
- the matrix $\boldsymbol{F}$ is generally non-quadratic and non-symmetric
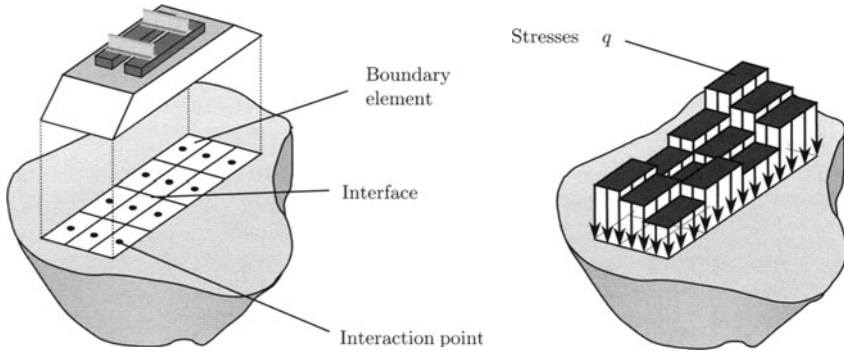
Fig. 2. Discretization of the interface assuming constant contact stresses

For the coupling with the finite element structure the flexibility matrix must be invertible, thus it must be quadratic. This can be accomplished by choosing identical shape functions for displacements and stresses. However, for the coupling the shape functions for displacements and stresses of finite elements should match the respective ones of boundary elements, but in commercial finite element software often elements with same order shape functions for displacements and stresses are not available. Hence, a compromise must be accepted.

Since ANSYS doesn't provide elements with independent shape functions for displacements and stresses we used finite elements with linear displacements and constant stresses. By choosing boundary elements with constant stresses and displacements, we satisfy the equilibrium, but we can enforce compatibility only at the midpoints (Fig. 2), which in our case become the interaction points.

Coupling of Structure and Subsoil

Inversion of (6) gives the contact stresses $\boldsymbol{q}$ as a function of the displacements $\boldsymbol{w}$

$$\boldsymbol{q} = \boldsymbol{F}^{-1} \cdot \boldsymbol{w} \ . \tag{7}$$

To satisfy equilibrium at the interface the corresponding quantities of the structure and the subsoil must have equal physical dimensions. Hence the contact stresses are transformed into nodal forces by a simple linear transformation

$$\boldsymbol{Q} = \boldsymbol{T}_q \cdot \boldsymbol{q} \tag{8}$$

with $\boldsymbol{T}_q$ a transformation matrix.

As mentioned above the compatibility of displacements is satisfied only at the interaction points. This relationship is given as

$$\boldsymbol{w} = \boldsymbol{T}_u \cdot \boldsymbol{u} \tag{9}$$

with $\boldsymbol{T}_u$ the corresponding transformation matrix. Plugging (7) and (9) into (8) we obtain

$$\boldsymbol{Q} = \underbrace{\boldsymbol{T}_q \cdot \boldsymbol{F}^{-1} \cdot \boldsymbol{T}_u}_{\boldsymbol{S}_S} \cdot \boldsymbol{u} \tag{10}$$

with $\boldsymbol{S}_S$ the complex stiffness matrix of the soil with respect to the interface nodes of the structure.

We finally plug (10) into the equation of motion (3b) and we obtain

$$(-\omega^2 \boldsymbol{M} + i\omega \boldsymbol{C} + \boldsymbol{K} + \boldsymbol{S}_S) \cdot \boldsymbol{u} = \boldsymbol{P} . \tag{11}$$

Finally, this linear equation system is solved using the ANSYS solvers.

Implementation in ANSYS

Since finite elements are well established in practice, very powerful FE packages are commercially available. They provide the full range of mechanical features like nonlinear material properties, large strains, and multi-physics, as well as the full range of numerical algorithms for nonlinearities, dynamic analysis in frequency and time domain, sparse solvers and iterative solvers, and also substructuring. Additionally they provide a software interface for user-supplied routines in nearly any stage of the solution process.

For this work we implemented the boundary element routines for the soil in ANSYS. Essentially we used the so-called User Programmable Features (UPF) and the substructuring features. The finite element model of the track is built using the ANSYS preprocessor. The soil is attached to the interface nodes by using the ANSYS element type MATRIX50, in terms of ANSYS a "superelement". The dynamic stiffness matrix of the superelement is read from an external file. Using UPF subroutines this file is overwritten with the respective matrix of the soil for each frequency step.

Application: Ballast Track and Slab Track System

For demonstration we show the results of a ballast track and a slab track system subjected to harmonic loads from 0–200 Hz. The geometry and FE-discretization can be seen in Fig. 3, whereas the unbounded halfspace is not shown, but included as described in the sections above. Both systems consist of two rails UIC60, rail pads, ties B70, and respectively a concrete slab over hydraulically stabilized base or a ballast layer. Attention should be paid to

the different stiffness of the rail pad, which is much lower for the slab track than for the ballast track. The rails are modelled with Timoshenko-beam elements, the pad with spring-damper-elements, and all other parts with 3D-solids. Damping mode is hysteretic except for the pad, which has viscous damping, and for the rail which has no damping at all. The loading are two harmonic vertical loads on the rail above tie zero.

All properties are taken from the benchmark test we contributed to, presented within this DFG priority programme [23]. Another presentation can be found in [14]. Hence, here we give only a brief overview of the mechanical properties in Tables 1 and 2.
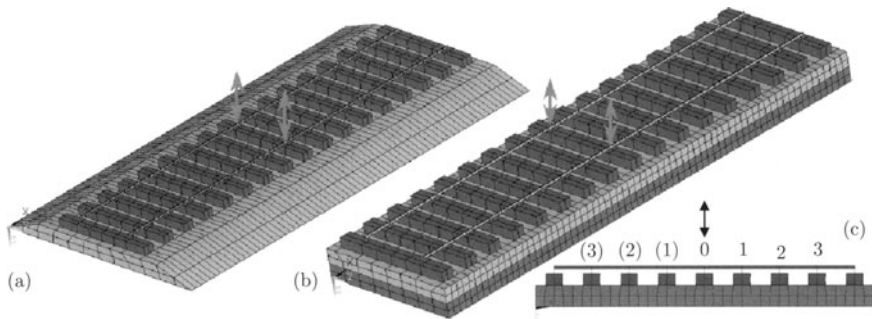
Fig. 3. Geometry and discretization of track systems, unbounded soil not shown. (a) Ballast track. (b) Slab track. (c) Tie numbering

Table 1. Material properties of the track components

	Thickness (m)	Shear wave velocity (m/s)	Mass density $(\mathrm{kg/m^3})$	Young's Modulus $(\mathrm{N/m^2})$	Poisson's ratio
Ballast	0.35	300	1700	3.8×10^8	0.25
Concrete Slab	0.30	2132	2500	3.0×10^{10}	0.25
Stabilized Base	0.30	1437	2200	1.2×10^{10}	0.16
Subsoil	∞	200	1700	1.7×10^8	0.16

Figure 4 shows the vertical flexibilities of both track systems with respect to the load position on the rail.

As can be seen the flexibilities are completely different, even the static rail-deflections above the loaded tie of the slab track are approximately three times larger than the respective deflections of the ballast track. The reason for this can be found in the properties of the pad. The ballast track pad is by

Table 2. Material properties of the rail pads

	Stiffness (N/m)	Damping (Ns/m)
Ballast track	6.00×10^8	2.0×10^4
Slab track	2.25×10^7	2.0×10^4
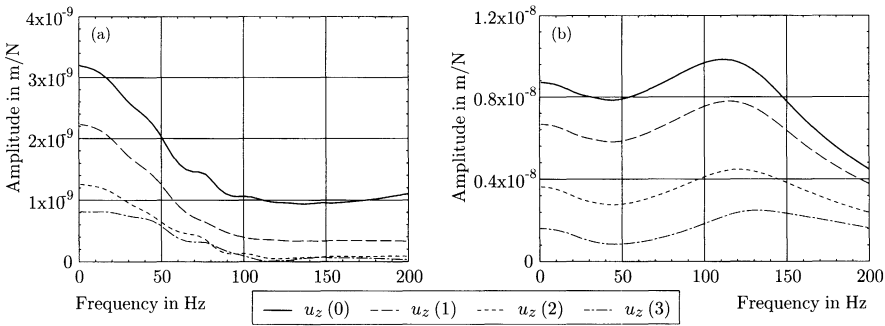
Fig. 4. Vertical flexibilities of the rail at tie 0–3. (a) Ballast track. (b) Slab track

more than one order stiffer than the slab track pad, therefore the deflections of the slab track highly correlate with the deflections of the pad, while the stiffness of the other parts have a relatively low influence. In contrast, the deflections of the ballast track are a result of the interaction of all components. In Fig. 5 we show the influence of different soil stratifications under a ballast track. The static response correlates with the stiffness of the halfspace. For low frequencies from 5 Hz to 50 Hz the response correlates more or less with the stiffness of the upper layer. For a soft layer over stiff halfspace even resonance frequencies emerge. For higher frequencies the response of the rail is getting more and more independent of the soil, but is rather determined by the properties of other track components.

For practical purposes not only the deflections of the track are important, but also the stresses in the track as well as the stresses in the soil. Figure 6 shows qualitatively the vertical normal stresses under a slab track subjected to a harmonic load of frequency 100 Hz. Due to the stiffness of the concrete slab and the stabilized base the contact stresses are uniformly distributed at the interface of base and soil, while for a ballast track the stresses are more concentrated under the two loads (not shown here).

Special Case: Distributed Surface Loads

Investigating ballast track systems, some authors need for the flexibility of the subsoil simplified solutions based on the response due to harmonic surface
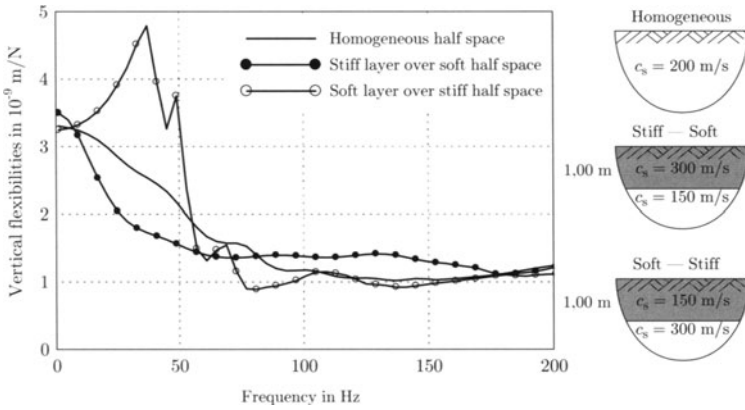
Fig. 5. Ballast track: vertical flexibilities of the rail at tie 0 for different soil stratifications
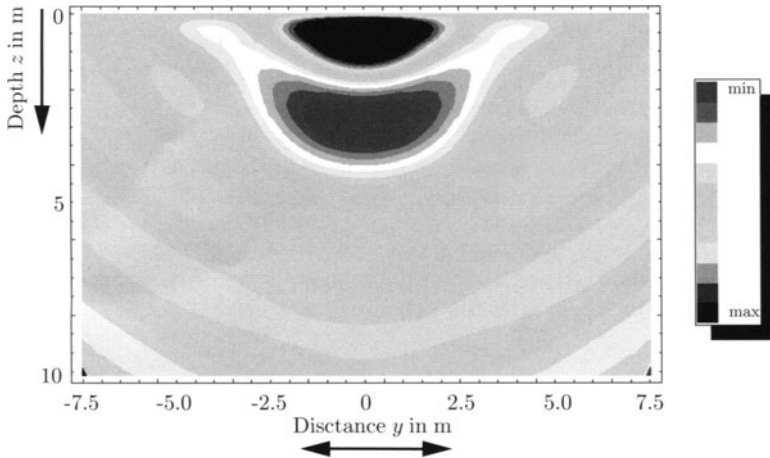
Fig. 6. Slab track: stresses σ_{zz} in the soil due to a harmonic load of 100 Hz

loads. The deflections and stresses of the soil due to that kind of load can be derived from the point load fundamental solutions described on p. 434 by means of numerical integration. This has been performed for vertical as well as for horizontal loads. The results were shared with the research groups of Professor Knothe (Berlin) and Professor Popp (Hannover), both participating in the DFG priority program, and used as input for their respective track models. A detailed discussion can be found in [26,27].

3.2 Analysis in Time Domain

The frequency domain analysis relies on the linearity of the structure and the soil, since to obtain the time history of deflections and stresses an Inverse

Fourier Transform (IFT) must be performed. But, the components of a track system, in particular soil, ballast, and pads, but also concrete slabs, exhibit nonlinear properties. Therefore the assumption of linear material properties is just an approximation which is valid only with limitations. To incorporate nonlinearities a time domain approach is necessary.

As in frequency domain analysis we use the substructure method in order to analyze track and subsoil separately.

Structure

The structure is again discretized by means of finite elements. Since we assume an arbitrary time dependence, we have to discretize spatially as well as in time. The equation of motion (2) is then given as

$$\boldsymbol{M} \cdot \ddot{\boldsymbol{u}}^i + \boldsymbol{C} \cdot \dot{\boldsymbol{u}}^i + \boldsymbol{K} \cdot \boldsymbol{u}^i = \boldsymbol{P}^i - \boldsymbol{Q}^i \ . \tag{12}$$

The vector $\boldsymbol{Q}^i$ denotes the resultant nodal forces of the contact stresses at the interface of track and soil and is a function of the displacements $\boldsymbol{u}^i$, while i is the time step index.

Subsoil

Comparing to frequency domain analysis the effort to get a fundamental solution in time domain is increasing, since the relationship of discretized stresses and displacements is now the convolution

$$\boldsymbol{w}(t) = \int\limits_0^t \boldsymbol{F}(t - \tau) \cdot \boldsymbol{q}(\tau) \, d\tau \tag{13}$$

with $\boldsymbol{w}$ the displacements of the interface nodes, $\boldsymbol{F}$ the soil flexibility given as the impulse response matrix of the layered halfspace, and $\boldsymbol{q}$ the stresses at the interface.

Unfortunately, a pure time domain solution for the flexibility matrix is only available for linear elastic homogeneous soil [7]. On the other hand, it is always possible to obtain a time domain solution from the frequency domain solution by means of Inverse Fourier Transform (IFT). Starting from (6)

$$\boldsymbol{w}(\omega) = \boldsymbol{F}(\omega) \cdot \boldsymbol{q}(\omega)$$

we obtain

$$\boldsymbol{w}(t) = \frac{1}{2\,\pi} \int\limits_{-\infty}^{\infty} \boldsymbol{F}(\omega) \cdot \boldsymbol{q}(\omega) \, e^{i\omega t} \, d\omega \ . \tag{14}$$

To obtain the impulse response matrix we subject the flexibility matrix in frequency domain to an IFT and we get

$$\boldsymbol{F}(t) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \boldsymbol{F}(\omega)\, e^{i\omega t}\, d\omega \; . \tag{15}$$

The numerical evaluation of (15) causes some trouble due to the high frequency contents of the flexibility matrix. Therefore, we use instead

$$\boldsymbol{F}(t) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \boldsymbol{F}(\omega)\, g(\omega)\, e^{i\omega t}\, d\omega \tag{16}$$

with $g(\omega)$ a Gaussian distribution as suggested in [28]

$$g(\omega) = \exp\left(-\sigma^2\omega^2\right) \; . \tag{17}$$

Here σ is the shape parameter. Equation (17) is the fourier transform of

$$g(t) = \frac{1}{2\,\sigma\sqrt{\pi}}\, \exp\left(-\frac{t^2}{4\,\sigma^2}\right) \; . \tag{18}$$

Using this so-called modified impulse response to evaluate the convolution (13) it leads to an approximation of the time-continuous stresses by a sequence of bell shaped impulses as illustrated in Fig. 7.
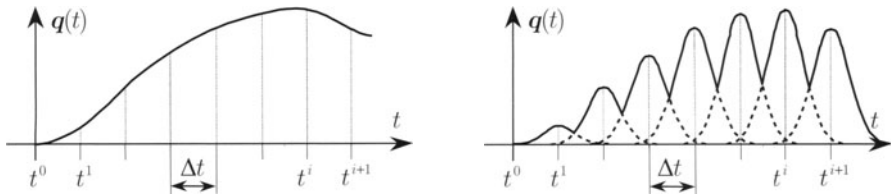
Fig. 7. Approximation of continuous contact stresses by a sequence of bell shape impulses

Now we discretize the convolution (13) in time and finally we get the relationship between stresses and displacements at the interface as

$$\boldsymbol{w}^{i+1} = \underbrace{\Delta t\, \boldsymbol{F}^i \cdot \boldsymbol{q}^1 + \cdots}_{\boldsymbol{w}^{\mathrm{hist}}} + \underbrace{\Delta t\, \boldsymbol{F}^0}_{\boldsymbol{F}^{\mathrm{cur}}} \cdot \boldsymbol{q}^{i+1}$$

$$\boldsymbol{w}^{i+1} = \boldsymbol{w}^{\mathrm{hist}} + \boldsymbol{F}^{\mathrm{cur}} \cdot \boldsymbol{q}^{i+1} \; . \tag{19}$$

Here $\boldsymbol{w}^{i+1}$ and $\boldsymbol{q}^{i+1}$ are the displacement vector and contact stress vector, respectively, at time step $i+1$, $\boldsymbol{F}^{\mathrm{cur}}$ is the current soil flexibility matrix, and $\boldsymbol{w}^{\mathrm{hist}}$ is the displacement vector due to contact stress history. We chose constant time steps to evaluate (19) efficiently, because in that case $\boldsymbol{F}^{\mathrm{cur}}$ doesn't change during calculation.

Coupling of Structure and Subsoil

Inversion of (19) yields

$$q^{i+1} = [F^{\mathrm{cur}}]^{-1} \cdot \left\{ w^{i+1} - w^{\mathrm{hist}} \right\} \; . \tag{20}$$

We satisfy compatibility

$$w^{i+1} = T_u \cdot u^{i+1} \tag{21}$$

and transform the stresses to nodal forces

$$Q^{i+1} = T_q \cdot q^{i+1} \tag{22}$$

with T_u and T_q the same transformation matrices already introduced in (8) and (9). Plugging (20) and (21) into (22) we obtain

$$Q^{i+1} = \underbrace{T_q \cdot [F^{\mathrm{cur}}]^{-1} \cdot T_u}_{K_S^{\mathrm{cur}}} \cdot u^{i+1} - \underbrace{T_q \cdot [F^{\mathrm{cur}}]^{-1} \cdot w^{\mathrm{hist}}}_{Q^{\mathrm{hist}}}$$

$$Q^{i+1} = K_S^{\mathrm{cur}} \cdot u^{i+1} - Q^{\mathrm{hist}} \; . \tag{23}$$

We plug (23) into the discrete equation of motion of the structure (12) to obtain

$$M \cdot \ddot{u}^{i+1} + C \cdot \dot{u}^{i+1} + [K + K_S^{\mathrm{cur}}] \cdot u^{i+1} = P^{i+1} + Q^{\mathrm{hist}} \; . \tag{24}$$

To solve (24) we use Newmark's β-method [4] with parameters for constant accelerations per time step. This scheme is implemented in ANSYS.

Partial Uplift

The formulation in time domain allows the treatment of nonlinear problems. As a first application we consider partial uplift at the interface in soil–structure interaction, which may occur in the case of ties on ballast.

The derivation above implies that each components of the track system is "welded" with its neighboring components. Therefore we also implemented an iteration scheme that takes into account a potential partial uplift. Details can be found in [14] and [25].

Application: Rail Pads with Nonlinear Material Properties

Rail pads generally don't have linear material properties, but are highly nonlinear. In this example we show the effect of the nonlinearities for a slab track system with the same dimensions as shown in Fig. 3, except that here only 13 ties were considered. The homogeneous soil has a shear modulus of $G = 90\,\mathrm{MN/m^2}$, a shear wave velocity of $c_{\mathrm{s}} = 230\,\mathrm{m/s}$, and a Poisson's ratio

of $\nu = 0.33$. The slab track is subjected to an axle load of 18 t moving with $v_{\text{Wheel}} = 100 \, \text{m/s}$. A rough estimation of the combined nonlinear stiffness of pad and rail mount is shown in Fig. 8.

The numbering of the ties is taken from Fig. 3. As one can see in Fig. 8 the deflections of the rail decreases considerably in case of nonlinear pad and mount.
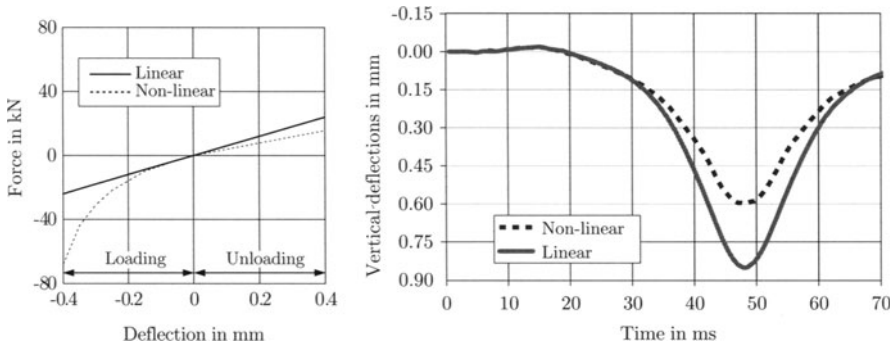
Fig. 8. Nonlinear properties of pad and mount. (a) Nonlinear combined spring characteristic. (b) Vertical deflections of rail above tie 2 for linear and nonlinear properties

Example: Simulation of a Passing Wheel Set

The following examples demonstrate the application of the time domain approach for moving loads. We used a simplified track model with two rails UIC 60 fixed on 50 ties B70 (Fig. 9). The ties are directly attached to homogeneous soil without any base or ballast in-between. The wheel set is modelled as two constant loads P moving with constant speed v_{Wheel}. The soil has a shear modulus of $G = 20 \, \text{MN/m}^2$, a shear wave velocity of $c_s = 100 \, \text{m/s}$, a Rayleigh wave velocity of $c_R = 335 \, \text{km/h}$, and a Poisson's ratio of $\nu = 0.33$.

Figure 10 shows the displacements of the 35th tie as a function of wheel speed. The results are scaled with the static deflections. As can be seen, the largest deformations occur when the wheel is moving with the Rayleigh wave velocity.

Figure 11 shows a zoomed in snapshot of the displacements of track and surrounding soil for a wheel set moving faster than Rayleigh wave velocity c_R. Clearly a Mach cone, well known from supersonic planes, and an uplift wave in front of the load has emerged.
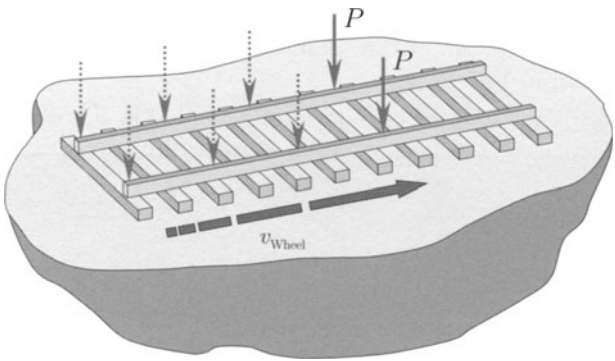
Fig. 9. Configuration of moving wheel set
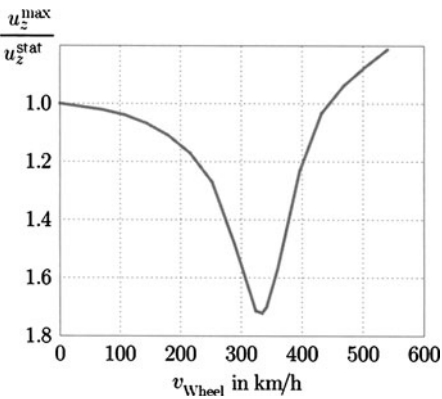
Fig. 10. Maximum vertical deflections of the 35th tie due to a moving wheel set

4   Embedded Structures – Analysis in Frequency Domain

The assumption of undisturbed and homogeneous soil below a track is usually not lifelike. Quite often the upper soil layer is replaced by bearable materials, sometimes fluids are injected, or even sand or gravel piles are built to increase the stiffness of the soil base. To take these inhomogeneities into account it is convenient to expand the finite element model of the track by an appropriate amount of the subgrade, which is equivalent to regarding track and surrounding soil as an embedded structure. As for structures on the surface we use the substructure method, but we have to take the excavation into account.

Structure

The structure with the surrounding soil is modelled by means of finite elements. There is no difference whatsoever concerning the governing equations, hence (4) is valid for embedded structures, too.
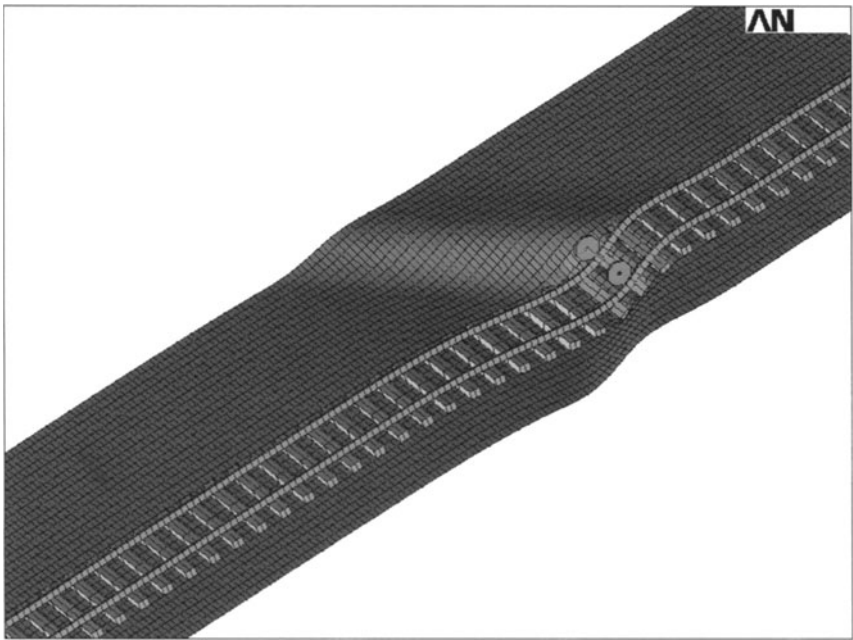
Fig. 11. Zoomed in napshot of displacements of track and soil for a moving wheel with "supersonic" speed ($v_{\text{Wheel}} = 400 \text{ km/h}$, $c_{\text{R}} = 335 \text{ km/h}$). The interior of the halfspace is not shown
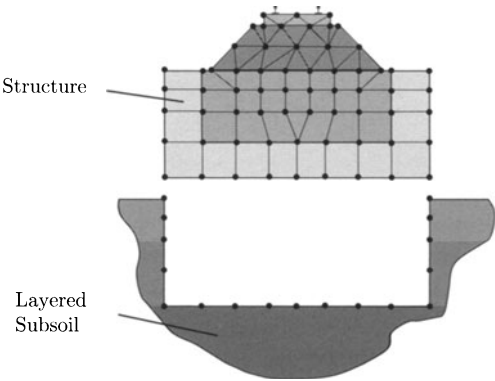
Fig. 12.   Embedded   structure with excavated soil

Subsoil

To obtain the flexibility matrix for the excavated halfspace with boundary elements at the interface only, we either have to compute fundamental solutions which incorporate the modified geometry of the excavated halfspace, or we have to modify the flexibility matrix of the regular halfspace. The former is rather tedious and would have to be repeated for every excavation shape. The latter is known from literature as "substructure deletion method" [11,22].
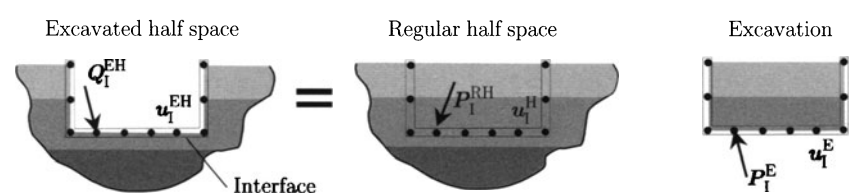
Fig. 13. Procedure of substructure deletion method

The method divides the soil into two substructures, the regular halfspace and the excavation. First we discretize the excavation using finite elements. The mass, damping, and stiffness matrix of the excavation are then condensed to the nodes at the interface.

The dynamic stiffness of the regular halfspace is computed as for surface structures, except that in case of embedded structures we have to evaluate the fundamental solutions also inside the soil. By satisfying equilibrium and compatibility conditions we obtain a dynamic stiffness matrix for the excavated halfspace essentially by subtracting the dynamic stiffness matrix of the excavation from the respective matrix of the regular halfspace [14] (Fig. 13).

We also implemented a second version of the substructure deletion method according to [13], which is restricted to embedded structures, while the version described above can also be applied for buried structures like tunnels.
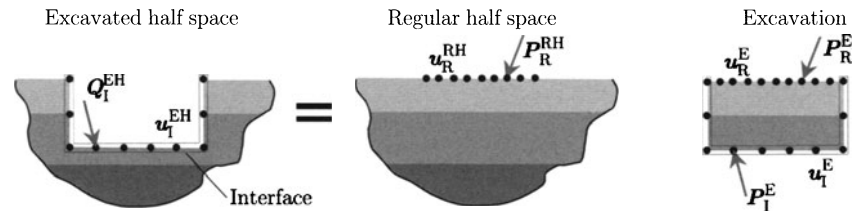
Fig. 14. Procedure of substructure deletion method according to [13]

The advantage of the second version is the fact that the fundamental solutions have to be evaluated only at the surface, therefore the effort is the same as for structures on the surface.

Coupling of Structure and Subsoil

The coupling procedure is identical to the coupling of surface structures. As soon as the stiffness matrix of the excavated halfspace is available we can compute a dynamic stiffness matrix of the complete system and solve the resulting linear equation system with ANSYS.

Application: Vibration Isolation by a Trench

A common measure to isolate vibrations triggered by rail and road traffic is a trench between track and neighboring buildings. The trench is either filled with a different material or is even empty. In the following we demonstrate the vibration isolation for a simplified model of that kind of situation, namely two massless rigid foundations on the surface separated by a trench filled with elastic material (15). The ratio of the shear modulus of the soil and the trench is $G_{\text{Trench}} = 0.001 \times G_{\text{Soil}}$ and the ratio of mass density is $\varrho_{\text{Trench}} = 0.55 \times \varrho_{\text{Soil}}$. The Poisson's ratio is $\nu = 0.45$.
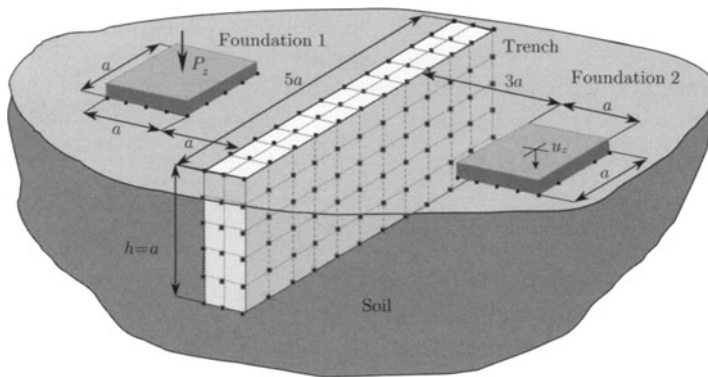
Fig. 15. Two rigid massless foundations separated by a trench

A harmonic vertical load P_z is applied to one foundation, while the vertical displacement u_z of the other is computed. As can be seen on the left hand side of Fig. 16 for static loads the trench has no effect whatsoever, while for dynamic loads a significant vibration reduction can be observed. On the right hand side the influence of the trench depth is shown. The deeper the trench the larger the amplitude reduction. For a depth of half the Rayleigh wave length the reduction in this example is approximately 50%.
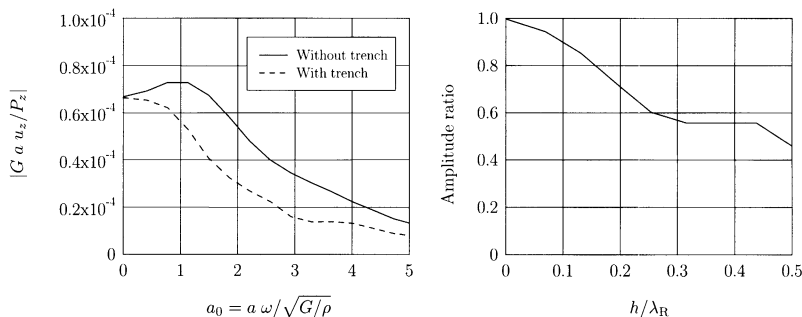
Fig. 16. Vibration isolation by a trench. (a) Modulus of normalized deflections of foundation 2 as a function of the dimensionless frequency a_0. (b) Ratio of vibration amplitude with trench vs. without trench as a function of the ratio of trench depth h and Rayleigh wave length λ_R

5   Outlook

The models described above can be used in the process of design and refitting of railway tracks. In particular the possibility to analyze the propagation of vibrations excited by moving trains and their isolation by constructional measures is substantial for railway lines in populated areas.

However, the models are by no means restricted to rail-bound traffic. Since we used the finite element method to analyze the structure, also vibrations induced by street-bound traffic and other man-made vibrations can be investigated, as well as vibrations due to seismic excitation.

Further research should be conducted to incorporate transient loads and non-linear material properties for embedded structures by means of a time domain approach.

6   Acknowledgement

References

1. Antes H., von Estorff O. (1989) Dynamic Response of Rigid Foundations and of Elastic Structures by Boundary Element Procedures. Soil Dyn Earthqu Eng Eng **8**:68–74
2. Apsel R.J., Luco J.E. (1987) Impedance Functions for Foundations Embedded in a Layered Medium: An Integral Equation Approach. Earthq Eng Struct Dyn **15**:213–231
3. de Barros F.C.P., Luco J.E. (1995) Dynamic Response of a Two-Dimensional Semi-Circular Foundation Embedded in a Layered Viscoelastic Half-Space. Soil Dyn Earthqu Eng **14**:45–57

4. Bathe K.J. (1982) Finite Element Procedures in Engineering Mechanics. Prentice-Hall, Englewood Cliffs

5. Betti R. (1997) Effects of the Dynamic Cross-Interaction in the Seismic Analysis of Multiple Embedded Foundations. Earthq Eng Struct Dyn **26**:1005–1019

6. Betti R., Abdel-Ghaffar A.G. (1994) Analysis of Embedded Foundations by Substructure Deletion Method. J Eng Mech **120**:1283–1303

7. Bode C. (2000) Numerische Verfahren zur Berechnung von Baugrund-Bauwerk-Interaktionen im Zeitbereich mittels Greenscher Funktionen für den Halbraum. Dissertation, Technische Universität Berlin

8. Bode C., Hirschauer R., Savidis S.A. (2002) Soil–structure interaction in the time domain using halfspace Green's functions. Soil Dyn Earthqu Eng **22(4)**:283–295

9. Bode C., Hirschauer R., Savidis S.A. (2000) Three Dimensional Time Domain Analysis of Moving Loads on Railway Tracks on Layered Soil. In: Schmid G., Chouw N. (Eds.) 2nd International Workshop on Wave Propagation, Moving Load, and Vibration Reduction, WAVE 2000, Bochum, Germany, December 13–15, 2000, AA Balkema, Rotterdam, 3–12

10. Brebbia C.A., Telles J.C.F., Wrobel L.C.(1984) Boundary Element Techniques. Springer-Verlag, Berlin

11. Dasgupta G. (1980) Foundation Impedance Matrices by Substructure Deletion. J Eng Mech **106**:517–524

12. von Estorff O., Kausel E. (1989) Coupling of Boundary and Finite Elements for Soil–Structure Interaction Problems. Earthq Eng Struct Dyn **18**:1065–1075

13. Gaul L. (1986) Machine–Foundation–Soil Interaction; Combined Continuum and Boundary Element Approach. RevBrMec **8**:169–198

14. Hirschauer R. (2001) Kopplung von Finiten Elementen mit Rand-Elementen zur Berechnung der dynamischen Baugrund-Bauwerk-Interaktion. Dissertation, Technische Universität Berlin

15. Hirschauer R., Bode C., Savidis S.A. (2001) Dynamik eingebetteter Bauwerke unter Berücksichtigung der Wellenabstrahlung. In: 19th CAD–FEM User's Meeting, October 17–19, Potsdam, Germany. CAD–FEM GmbH, Grafing, Chapter 1.5.11

16. Hirschauer R., Bode C., Savidis S.A., Haaß R. (2000) Dynamische Baugrund-Bauwerk Wechselwirkung mit ANSYS unter Berücksichtigung der Wellenabstrahlung – Lösung im Zeitbereich. In: 18th CAD–FEM User's Meeting, September 20–22, Friedrichshafen, Germany. CAD–FEM GmbH, Grafing, Chapter 1.2.1

17. Hirschauer R., Bode C., Savidis S.A. (1999) Frequenzbereichs-Analyse der dynamischen Baugrund-Bauwerk Wechselwirkung mit ANSYS unter vollständiger Berücksichtigung der Wellenabstrahlung. In: 17th CAD–FEM User's Meeting, Sonthofen, Germany. CAD–FEM GmbH, Grafing

18. Kausel E. (1994) Thin-Layer Method: Formulation in the Time-Domain. Int J Numer Meth Engng **37**:927–941

19. Kausel E. (1986) Wave Propagation in Anisotropic Layered Media. Int J Numer Meth Engng **23**:1567-1578

20. Lamb H. (1904) On the Propagation of Tremors Over the Surface of an Elastic Solid. Philos T Roy Soc **203**:1-42

21. Mita A., Luco J.E. (1987) Dynamic Response of a Square Foundation Embedded in an Elastic Half-Space. Comput methods appl mech eng **63**:233–259

22. Neto E.M., Romanini E., de Pontes B.R. (1995) A Boundary Element Implementation of the Substructure Deletion Method. In: Boundary Elements XVII. Computational Mechanics Publications, Southampton, 375–386

23. Rücker W., et al. (2002) A comparative study of results from numerical track–subsoil calculations. DFG–Priority–Programme "System Dynamics and Long-Term Behaviour of Vehicle, Track, and Subgrade". To be enclosed in this volume.

24. Savidis S.A., Bode C., Hirschauer R. (1999) 3–Dimensionale dynamische Baugrund–Bauwerk Wechselwirkung infolge seismischer Erregung unter Berücksichtigung klaffender Fugen. In: Savidis S.A. (Ed.) Entwicklungsstand in Forschung und Praxis auf den Gebieten des Erdbebeningenieurwesens, der Boden- und Baudynamik: Vortragsband der Dreiländertagung D-A-CH 1999 der Deutschen Gesellschaft für Erdbeben–Ingenieurwesen und Baudynamik, Berlin, Germany, Nov. 24–25, 1999. DGEB, Berlin

25. Savidis S.A., Bode C., Hirschauer R., Hornig J. (1999) Dynamic Soil-Structure Interaction with Partial Uplift. In: Frýba L., Náprstek J. (Eds.) Structural Dynamics – EURODYN '99. A.A. Balkema, Rotterdam Brookfield, 957-962

26. Savidis S.A., Hirschauer R. (1997) Dynamische Steifigkeiten von Schwellensystemen auf geschichtetem Untergrund. Erster Zwischenbericht zum DFG-Schwerpunktprogramm "Systemdynamik und Langzeitverhalten von Fahrwerk, Gleis und Untergrund", Fachgebiet Grundbau und Bodenmechanik der Technischen Universität Berlin

27. Savidis S.A., Hirschauer R. (1999) Dynamische Steifigkeiten von Schwellensystemen auf geschichtetem Untergrund. Zweiter Zwischenbericht zum DFG-Schwerpunktprogramm "Systemdynamik und Langzeitverhalten von Fahrwerk, Gleis und Untergrund", Fachgebiet Grundbau und Bodenmechanik der Technischen Universität Berlin

28. Wolf J.P. (1988) Soil–Structure–Interaction Analysis in Time Domain. Prentice Hall, Englewood Cliffs

Rigid Body Dynamics of Railway Ballast

Thomas Schwager and Thorsten Pöschel

Humboldt-Universität zu Berlin, Charité, Institut für Biochemie,
Monbijoustraße 2, D-10117 Berlin, Germany

Abstract. A method for the discrete particle simulation of almost rigid, sharply edged frictional particles, such as railway ballast is proposed. In difference to Molecular Dynamics algorithms, the method does not require knowledge about the deformation-force law of the material. Moreover, the method does not suffer from numerical instability which is encountered in MD simulations of very stiff particles.

1 Introduction

Traditionally, the subgrade of railway tracks is modeled using continuum mechanics. These methods have been proven to yield reliable results in many applications and have been developed to standard methods. In certain applications, however, continuum models fail in describing the mechanical properties of the subgrade. This is the case when the ballast must not be considered as a continuous medium, but when the granularity of the material is important. Typical processes which cannot be sufficiently explained by continuum models are sedimentation of the ballast due to recurrent load, abrasion of the ballast particles which leads to less efficient damping properties and the formation of force chains inside the ballast material.

In the past decade mainly by physicists much work has been done in the field of Molecular Dynamics of granular material, i.e., the numerical simulation of granular material as many-particle systems. This technique was applied to many interesting systems and has contributed to the explanation of several exciting and technologically important effects, such as mixing and demixing of granular materials, avalanche statistics on sand heaps, milling processes, convection dynamics in shaken granular materials and others. Many examples of such simulations can be found, e.g., in [1].

The idea of Molecular Dynamics is to simulate the granular material as a many particle system and to determine the dynamics of the system by numerical integration of Newton's equation of motion for each of the N particles:

$$
\begin{aligned}
\ddot{\boldsymbol{r}}_i &= \boldsymbol{F}_i / m_i \\
\ddot{\boldsymbol{\phi}}_i &= \hat{J}_i^{-1} \boldsymbol{M}_i \,,
\end{aligned}
\tag{1}
$$

where $\boldsymbol{r}_i$ and $\boldsymbol{\phi}_i$ are the position and the orientation of the i-th particle of mass m_i and moment of inertia $\hat{J}_i$ while $\boldsymbol{F}_i$ and $\boldsymbol{M}_i$ are the force and the torque acting on this grain. In three dimensions Eqs. (1) establish a set of

$6N$ coupled non-linear differential equations which have to be numerically integrated. The force $\boldsymbol{F}_i$ consists of gravity and of the interaction force of the particle i with other particles j

$$\boldsymbol{F}_i = m_i \boldsymbol{g} + \sum_j \boldsymbol{F}_{ij}\,, \tag{2}$$

where $\boldsymbol{g}$ is the gravitational acceleration. Granular particles interact with each other only if they are in mechanical contact. For spheres of radii R_i and R_j we write

$$\boldsymbol{F}_{ij} = \begin{cases} F_{ij}^N \boldsymbol{n}_{ij} + F^T \boldsymbol{t} & \text{if } |\boldsymbol{r}_i - \boldsymbol{r}_j| < R_i + R_j \\ 0 & \text{else}, \end{cases} \tag{3}$$

where F^N and F^T are the components of the force in normal and tangential direction with respect to the inter-center vector $\boldsymbol{r}_i - \boldsymbol{r}_j$ and $\boldsymbol{n}$ and $\boldsymbol{t}$ are the corresponding unit vectors. There exist several models for the interaction forces in normal and tangential direction $F_{ij}^{N/T}\left(\boldsymbol{r}_i, \boldsymbol{r}_j, \dot{\boldsymbol{r}}_i, \dot{\boldsymbol{r}}_j\right)$ (see, e.g. [2]) which shall not be discussed in detail here.

2   Molecular Dynamics Fails for the Simulation of Railway Ballast

There are many examples where granular systems have been simulated by Molecular Dynamics, however, except for few examples, realistic simulations have been achieved only for systems where the *dynamical* behavior of the grains dominates the system properties. When the static properties of the particle system become important, i.e., when the relative velocities are small or zero, the details of the interaction force become essential for understanding the system behavior. We are faced with two major problems:

- The interaction force of contacting particles must be known as a function of the particle positions and velocities. In the case of sharply edged grains such as railway ballast, this function is unknown.
- As soon as the realistic simulation of static properties matters for the system behavior, the simulation slows down extremely which implies that to achieve affordable computation times one has to make simplifying assumptions on the particle contact which are not justified from the point of view of mechanics and material science.

For these reasons we believe that the described Molecular Dynamics method is *in principle* unsuitable for the simulation of railway ballast. Below we list arguments which support this statement:

1. **Elastic normal force:** Whereas the normal elastic force of contacting spheres is given by the Hertz law

$$F_{ij}^{N,\text{el}} = \frac{2}{3} \frac{Y \sqrt{R_{\text{eff}}}}{1 - \nu^2} \, \xi_{ij}^{3/2} \tag{4}$$

with Y, ν, and R_{eff} being the Young modulus, the Poisson ratio, and the effective radius, respectively, this force is not known for more complicatedly shaped particles. For smooth particles (when the local radius of the contact area is large as compared to the compression $\xi_{ij} \equiv R_i + R_j - |r_i - r_j|$) the function (4) is certainly a good approximation, for sharply edged particles, however, this function fails. Some authors assume that the normal force is proportional to what they call "overlap", i.e., the volume of the compressed material or (yet more simple) in two dimensions the "compression length" (for spheres the value ξ) which is certainly incorrect even for the most simple case of contacting spheres and the more for more complicatedly shaped grains.

2. **Dissipative normal force:** The dissipative normal force of contacting bodies is unknown, in general. For viscoelastically colliding spheres and other smooth contacting surfaces it is given by

$$F_{ij}^{N,\text{diss}} = A\dot{\xi}\frac{\partial}{\partial \xi}F^{N,\text{el}}. \tag{5}$$

The prefactor A is a complicated function (for details see [3]) which contains the viscous constants of the material which are unknown in general. For particles which are not smooth, such as railway ballast, even the functional form of the dissipative force is unknown. The frequently applied force law $F_{ij}^{N,\text{diss}} \propto \dot{\xi}$ is not justified and fails even for spheres.

3. **Tangential force:** Whereas for smooth bodies the normal force can be determined from bulk properties of the material, the tangential force is determined by the bulk and by surface properties. A natural (phenomenological) assumption is

$$F_{ij}^{T} \leq \mu F_{ij}^{N}, \tag{6}$$

where μ is the Coulomb friction parameter. Unfortunately, this model is not sufficient for static systems: Assume a (non-spherical) particle which rests on an inclined plane (angle α). Its normal force is $F^N = mg\cos\alpha$ and there is a corresponding tangential force too. To prevent the particle from sliding along the plane one has to assume an additional force which mimics static friction, e.g. [2]. This force cannot be derived from material or surface properties, hence, it is arbitrary. Choices for this function which can be found in literatures are even in disagreement with basic mechanics.

4. **Numerical integration of Newton's law:** Besides the fact that the numerical integration of the set (1) for a relevant number of particles N, e.g., $N = 1,000$ which establish a quite small system of $10 \times 10 \times 10$ particles in three dimensions, requires serious numerical effort, there are principal problems when applying Molecular Dynamics to a particle system:

- For very rigid particles the gradient of the force is large and, therefore, requires a very small integration time step. The more rigid the particle

material the slower progresses the simulation. Assuming a time step of $\Delta t = 10^{-7}$ sec which is used in many MD simulations of granular material, one needs 10^9 integration steps of the equations of motion (1) to achieve a real time of 100 sec only. Hence, a simulation of long time behavior using MD seems to be unrealistic.

- The accuracy of the algorithm cannot be increased too much by reducing the time step. Frequently, predictor-corrector algorithms are used, which require powers of the time step, $(\Delta t)^2 = 10^{-14}$, $(\Delta t)^3 = 10^{-21}$, etc. (again for $\Delta t = 10^{-7}$). If we ignore the prefactor of this powers in the integration scheme to obtain a crude estimate, one has to sum numbers of the orders $O(1)$, $\mathcal{O}\left(10^{-7}\right)$, $\mathcal{O}\left(10^{-14}\right)$, etc. Double precision numbers have a mantissa of 15 digits, therefore, one cannot sum numbers which are different by more than 10^{15}. This limits the minimal value of the integration time step. If one requires higher precision one needs real numbers in quadruple precision or higher which implies higher memory consumption and (more importantly) yet larger computation time since multiplication of two numbers in quadruple precision requires approximately quadruple time as a multiplication of double precision numbers.

We are aware that these numbers are a very crude estimate which have been given to illustrate the problem. Of course, more sophisticated integration schemes suffer less critical from the addition of different numbers. In principle, however, the problem persists.

3   Rigid Body Dynamics

In Molecular Dynamics simulations the trajectories of the particles are determined by numerical integration of Newton's equations of motion. As discussed above, this implies the "soft particle assumption", i.e., the particle deform under load. Since the particles are very hard but not completely rigid we encounter serious numerical problems as described above.

The Rigid Body Dynamics originates from the opposite idea: The interaction forces are determined from the required behavior of the particles. This method is, therefore, suited to simulate perfectly rigid particles without the necessity to specify a certain force-deformation law. We consider the example of a rigid sphere which rests on a rigid flat surface (see Fig. 1). There is a point-like contact of the sphere and the surface, hence, there is a contact force F^N in vertical direction. Moreover gravity acts on the sphere causing a force $-mg$. If one would chose the contact force $F^N = 0$, the sphere would move downwards with an acceleration g, i.e., it would penetrate the surface. This unphysical behavior has to be avoided by the proper choice of the force F^N due to the following conditions:

1. *Contact forces have to be chosen in order to avoid penetration of contacting particles.*
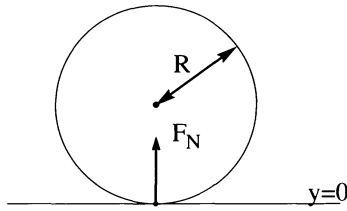
Fig. 1. A rigid sphere resting on a rigid flat surface.

For our example $F^N \geq mg$ follows. On the other hand, a force $F^N > mg$ would lead to unphysical behavior since the sphere would move upwards. This is avoided by the second condition.

2. *A contact force vanishes when the contact breaks.*
   A contact is said to break if the normal component of the relative acceleration of the concerned particles, or their normal velocity is larger than zero. (The relative velocity is counted positive if the particles separate from each other.)

3. *There are no attractive normal forces.*
   In our example $F^N > mg$ causes the contact to break which implies that the contact force vanishes. Therefore, from these conditions follows $F^N = mg$. For this choice the total force is zero and the sphere rests on the plane. The conditions 1.-3. are sufficient to describe any particle system as long as there are no friction forces. For systems with friction we need one more condition:

4. *Friction forces act in parallel with the contact plane, i.e., perpendicular to F^N. Given the tangential force F^* which is necessary to keep two particles from sliding. Then the acting tangential force is $|F| = \min\left(|F^*|, |\mu F^N|\right)$. Its sign has to be chosen opposite to the tangential relative acceleration (or the tangential relative velocity).*
   In agreement with Coulomb's friction law the particles slide only if $|F^T| \geq |F^N|$. If this condition is fulfilled the friction force adopts its maximal value $\pm \mu F^N$.

To perform simulations one has to derive the forces of the particles in normal and tangential directions from these four conditions. The corresponding algorithm will be explained in sections 5 and 6.

In our simple example (Fig. 1) there exists only one contact between the sphere and the plane. In more complex situations, e.g. for a resting cube, there are contacts areas instead of points. These contacts may be always reduced to point contacts. It will be shown that the described conditions are sufficient to determine the forces and torques which act on the particles, provided there are not too many contacts in the system. If the number of contacts is too large only the total force and the total torque which act on a particle may be determined, but not each of the pairwise contact forces.

For the computation of the particle trajectories, however, the total forces and torques are sufficient. We will return to this issue in Sec. 8.

Rigid Body Dynamics has been intensively studied in the past two decades. Descriptions of the algorithm can be found in [4,5]. The core of the algorithm is the numerical computation of the contact forces which is a Linear Complementarity Problem [6,7]. An efficient algorithm for this type of problems can be found in [8–11]. Rigid Body Dynamics has also been applied to granular systems, e.g. [12,13], where frictionless smooth spheres have been simulated. Systems of granular particles subject to friction have been studied, e.g., in [14].

Due to our understanding the Rigid Body Dynamics is much better suited for the simulation of railway ballast for the following reasons:

- Ballast particles are irregularly shaped and sharply edged. Even if the bulk material properties were precisely known, the contact force law is unknown due to the complicated shape.
- Ballast particles are very stiff which implies that the gradient of the interaction force is very steep. In this regime the numerical integration of Newton's equation is problematic. For the dynamics of the system the deformation of single particles is unimportant, i.e., the Rigid Body assumption is well justified.
- Static friction, whose treatment in MD-simulations is problematic too, is essential for the dynamics of the system. It is correctly modeled in Rigid Body Dynamics.
- The long time behavior of railway ballast is affected by abrasion and fragmentation of particles. Molecular Dynamics of fragmenting particles may cause artifacts for several reasons which cannot be discussed here in detail (see [15]). Rigid Body Dynamics is very well suited for this case.

4   Schedule of Rigid Body Simulations

The state of the granular system is described by the position and orientation of its particles and by the according time derivatives. Contacts between the particles may be classified as sticking and sliding contacts. In the due of time the contact network is modified by creation and termination of contacts as well as by transformation of sticking contacts into sliding ones and vice versa. Whenever the contact network is modified the state of the system changes qualitatively. The simulation proceeds in discrete time steps. Each of them consists of

1. **Contact detection:** all existing contacts are registered.
2. **Treatment of collisions:** A collision takes place if two contacting particles move with negative normal relative velocity. In this case one cannot determine a *finite* contact force which would avoid penetration of the particles since *any* force, however large it is, would need a short but finite

path to retard the colliding particles. Hence, mutual penetration would be unavoidable. Therefore, we need a special treatment for collisions (see Sec. 7).

3. **Formulation of the geometry equations:** After a collision, in general, there is a number of contacts of particles which have a positive normal relative velocity, i.e., the particles lose contact. These contacts have to be erased from the list of contacts. The normal components of the relative velocities at all remaining contacts are zero. We have to establish the geometry equations which contain the information about the geometry of the system (see Sec. 5).

4. **Computation of the forces:** Section 6 deals with the computation of the relative accelerations by means of the geometry equations.

5. **Integration of the equations of motion:** Finally we have to integrate the equations of motion for all particles. During this operation it may be necessary to update the geometry equations and to repeat the computation of the forces according to the integration scheme used.

5   Mathematical Description

For convenience at first we will restrict to frictionless particles. When the mathematical framework has been developed for this simplified case we will then introduce friction forces between particles. The rigidity of the particles is enforced by means of mathematical motion constraints of the form

$$g(q) \geq 0 \ , \tag{7}$$

where q is a vector containing the positions (center of mass position and orientation) of all particles of the system. The constraint function g shall be zero if particles in the system are in contact and larger than zero otherwise. For spheres, which is the most simple case, the constraint function reads

$$g(q) = |\boldsymbol{r}_i - \boldsymbol{r}_j| - R_i - R_j \ . \tag{8}$$

If the spheres would deform each other ($|\boldsymbol{r}_i - \boldsymbol{r}_j| < R_i + R_j$) the function $g(q)$ would be negative, if the particles touch each other it would be zero. To prevent the deformation we require $g(q)$ to be positive or zero, i.e., it has to fulfill the condition (7). For sharp edged particles, such as particles which are described by polyhedrons or polygons, the motion constraints are

$$g(q) = \boldsymbol{n}_j \left(\boldsymbol{r}_i + \boldsymbol{x}_i - \boldsymbol{r}_j \right) - d \ , \tag{9}$$

where $\boldsymbol{n}_j$ is normal of the face of particle j which is in contact with an edge of the particle i. The vectors $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$ are the center of mass positions of the particles, the edge of particle i that is in contact with the particle j is at position $\boldsymbol{r}_i + \boldsymbol{x}_i$. The constant d is the distance of the contacting face of j
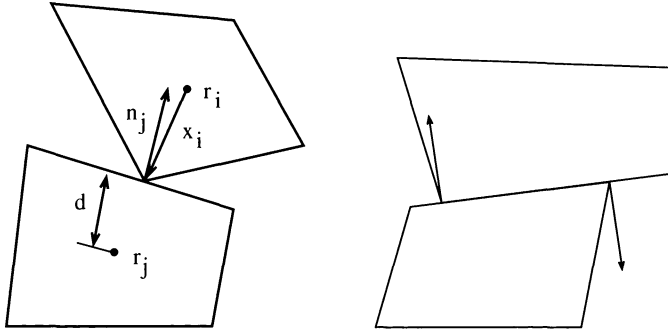
Fig. 2. Left: a face-edge contact. r_i and r_j are the center of mass positions, x_i is the coordinate of the contacting edge relative to the center of mass of particle i and n_j is the normal of the contacting face of particle j. Right: face-face contacts can be reduced to two face-edge contacts. The face normals of each of the contacts are displayed as well.

from the center of the particle. The left picture in Fig. 2 shows a sketch of a typical contact of two particles. Face-face contacts can be described by two face-edge contacts (see Fig. 2).

Every motion constraint corresponds to a scalar contact force f. According to d'Alembert's principle the direction of the contact force is given by the spacial derivative of g with respect to all components of the coordinate vector q, namely $\partial g / \partial q$. The contact force that acts on a certain particle i is $f \partial g / \partial q_i$, with q_i being the coordinates of particle i. We can formulate the equation of motion for the particles[1]

$$\hat{M}_i \ddot{q}_i = Q_i + \sum_\alpha f_\alpha \frac{\partial g_\alpha}{\partial q_i} \tag{10}$$

$$g_\alpha(q) \geq 0 \ . \tag{11}$$

$\hat{M}_i$ is the mass matrix of the particle i, which has the form

$$\hat{M}_i = \begin{pmatrix} m_i & 0 & 0 & 0 \\ 0 & m_i & 0 & 0 \\ 0 & 0 & m_i & 0 \\ 0 & 0 & 0 & \hat{J}_i \end{pmatrix} \ , \tag{12}$$

where $\hat{J}_i$ is the moment of inertia tensor. In two dimensional systems $\hat{J}_i$ is only a scalar J_i and there are only 2 entries of m_i. Q_i finally is the external force (and torque) acting on particle i. This is usually gravitation, but other external forces can be incorporated at this point as well.

Although we formulated the motion constraints in the form $g(q) \geq 0$ to allow separation of particles, contact forces can only act if particles are actually in

---

[1] Particle indices are written in Latin letters, contact indices in Greek letters.

contact. Therefore, constraints which are strictly positive, i.e. $g(q) > 0$ (the particles are separated), can be disregarded. These constraints are said to be inactive. The remaining constraints, the active ones, are thus satisfied by $g(q) = 0$. Since the $g(q)$ have to remain non-negative their time derivatives have to be non-negative as well. We therefore have

$$\dot{g}_\alpha = \frac{\partial g_\alpha}{\partial q_k} \dot{q}_k \geq 0 \tag{13}$$

$$\ddot{g}_\alpha = \frac{\partial g_\alpha}{\partial q_k} \ddot{q}_k + \frac{\partial^2 g_\alpha}{\partial q_k \partial q_l} \dot{q}_k \dot{q}_l \geq 0 . \tag{14}$$

For simplicity of notation we used the Einstein convention, i.e. summation over doubly occurring indices k and l is implied. These time derivatives are the relative velocity and relative acceleration of the particles at their contact points. It is important to note that $\dot{g}_\alpha$ and $\ddot{g}_\alpha$ are not the relative velocity or acceleration of the centers of mass of the particles but of the points of both particles which are actually in contact. It can easily happen that the relative velocity or acceleration of the contact points are positive (the particles are about to separate) although their centers of mass approach each other.
We insert the equation of motion (10) into (14) and find

$$\ddot{g}_\alpha = \frac{\partial g_\alpha}{\partial q_k} \left[\hat{M}_k^{-1} \left(Q_k + \sum_\beta f_\beta \frac{\partial g_\beta}{\partial q_k} \right) \right] + \frac{\partial^2 g_\alpha}{\partial q_k \partial q_l} \dot{q}_k \dot{q}_l$$

$$= \frac{\partial g_\alpha}{\partial q_k} \left(\hat{M}_k^{-1} Q_k \right) + \frac{\partial^2 g_\alpha}{\partial q_k \partial q_l} \dot{q}_k \dot{q}_l + \frac{\partial g_\alpha}{\partial q_k} \hat{M}_k^{-1} \left(\sum_\beta f_\beta \frac{\partial g_\beta}{\partial q_k} \right) . \tag{15}$$

The first term on the right hand side describes the action of the external forces, the second term describes the action of inertial forces as, e.g., centrifugal force and Coriolis force, the third term finally describes the action of the contact forces. We can rewrite this equation as

$$\ddot{g}_\alpha = b_\alpha + \sum_\beta A_{\alpha\beta} f_\beta , \tag{16}$$

with $A_{\alpha\beta}$ and b_α abbreviating

$$A_{\alpha\beta} = \frac{\partial g_\alpha}{\partial q_k} \hat{M}_k^{-1} \frac{\partial g_\beta}{\partial q_k}$$

$$b_\alpha = \frac{\partial g_\alpha}{\partial q_k} \left(\hat{M}_k^{-1} Q_k \right) + \frac{\partial^2 g_\alpha}{\partial q_k \partial q_l} \dot{q}_k \dot{q}_l . \tag{17}$$

From now on we will denote the relative acceleration of the contacting particles at their contact points – the contact acceleration – as a_α instead of $\ddot{g}_\alpha$. Equation (16) turns into

$$a_\alpha = b_\alpha + A_{\alpha\beta} f_\beta , \tag{18}$$

where again summation over β is implied. We will call this equation the geometry equation. By means of this equation and the consistency conditions introduced in Sec. 3 we can now determine the contact forces f_β. The consistency conditions read

$$a_\alpha \geq 0$$
$$f_\alpha \geq 0 \tag{19}$$
$$a_\alpha f_\alpha = 0 \ .$$

The first condition prevents deformation of particles, the second one excludes attractive forces and the third one requests that contact forces may only act if the particles stay in contact, i.e., if $a_\alpha = 0$. These conditions together with the geometry equation (18) allows to determine the unknown contact forces f_α. The whole system consisting of Eq. (18) and the conditions (19) is called a Linear Complementarity Problem. It can be solved by Dantzig's algorithm [6].

To incorporate friction we introduce additional motion constraints which shall, if possible, impede a tangential motion of the contacting particles. For polygonal particles they are of the form

$$g(q) = t_j \left(r_i + x_i - r_j - x_j \right) \ . \tag{20}$$

This constraint has a similar form as the constraint of the normal motion of the particles (9) but instead of the normal unit vector of the contacting face of particle j the tangential unit vector t_j appears, thus ensuring that the edge of particle i at position $r_i + x_i$ does not move along the face of particle i away from the point $r_j + x_j$ of initial contact.

These motion constraints are, however, of different nature than the normal motion constraints. Whereas in the case of the normal motion the constraints must never be violated, the constraints on the tangential motion may actually be violated, as it happens when the particles start to slide. This is due to the fact that the magnitude of the friction forces are limited by μf_N, with μ being the friction constant and f_N the corresponding normal force. As reflected by the consistency condition 4 (ref. Sec. 3) the friction force must adopt its maximum value if particles actually slide, i.e. if the tangential motion constraint is inactive. Thus, in this case the value of the friction force is determined without need of further consideration. In contrast to the case of normal motion constraints we may not neglect the inactive constraints because now the corresponding contact forces are non-zero and thus the constraint has to be kept in consideration in order to determine the direction of the tangential force.

Since friction causes only further motion constraints there is, in principle, no need of further discussion of the problem. The geometry equation (18) can describe systems with friction as well. For simplicity of notation it is worth, however, to consider normal and friction forces and their corresponding mo-

tion constraints separately. We now have 2 classes of motion constraints

$$g^N(q) \geq 0$$
$$g^T(q) = 0 \qquad (21)$$

and the corresponding contact forces f^N and f^T. Now the equation of motion reads

$$\hat{M}_k \ddot{q}_k = Q_k + \sum_\alpha \left(f_\alpha^N \frac{\partial g_\alpha^N}{\partial q_k} + f_\alpha^T \frac{\partial g_\alpha^T}{\partial q_k} \right) . \qquad (22)$$

For the second time derivative of the motion constraints we obtain

$$\ddot{g}_\alpha^N = \frac{\partial g_\alpha^N}{\partial q_k} \left(\hat{M}_k^{-1} Q_k \right) + \frac{\partial^2 g_\alpha^N}{\partial q_k \partial q_l} \dot{q}_k \dot{q}_l + \frac{\partial g_\alpha^N}{\partial q_k} \hat{M}_k^{-1} \left(\sum_\beta f_\beta^N \frac{\partial g_\beta^N}{\partial q_k} + f_\beta^T \frac{\partial g_\beta^T}{\partial q_k} \right)$$

$$\ddot{g}_\alpha^T = \frac{\partial g_\alpha^T}{\partial q_k} \left(\hat{M}_k^{-1} Q_k \right) + \frac{\partial^2 g_\alpha^T}{\partial q_k \partial q_l} \dot{q}_k \dot{q}_l + \frac{\partial g_\alpha^T}{\partial q_k} \hat{M}_k^{-1} \left(\sum_\beta f_\beta^N \frac{\partial g_\beta^N}{\partial q_k} + f_\beta^T \frac{\partial g_\beta^T}{\partial q_k} \right) . \qquad (23)$$

Renaming again $\ddot{g}_\alpha^{\{N,T\}}$ by $a_\alpha^{\{N,T\}}$ and using the abbreviations

$$b_\alpha^N = \frac{\partial g_\alpha^N}{\partial q_k} \left(\hat{M}_k^{-1} Q_k \right) + \frac{\partial^2 g_\alpha^N}{\partial q_k \partial q_l} \dot{q}_k \dot{q}_l$$

$$b_\alpha^T = \frac{\partial g_\alpha^T}{\partial q_k} \left(\hat{M}_k^{-1} Q_k \right) + \frac{\partial^2 g_\alpha^T}{\partial q_k \partial q_l} \dot{q}_k \dot{q}_l \qquad (24)$$

and

$$A_{\alpha\beta}^{NN} = \frac{\partial g_\alpha^N}{\partial q_k} \hat{M}_k^{-1} \frac{\partial g_\beta^N}{\partial q_k} \qquad A_{\alpha\beta}^{NT} = \frac{\partial g_\alpha^N}{\partial q_k} \hat{M}_k^{-1} \frac{\partial g_\beta^T}{\partial q_k}$$

$$A_{\alpha\beta}^{TN} = \frac{\partial g_\alpha^T}{\partial q_k} \hat{M}_k^{-1} \frac{\partial g_\beta^N}{\partial q_k} \qquad A_{\alpha\beta}^{TT} = \frac{\partial g_\alpha^T}{\partial q_k} \hat{M}_k^{-1} \frac{\partial g_\beta^T}{\partial q_k} \qquad (25)$$

we can write the modified geometry equations

$$a_\alpha^N = b_\alpha^N + \sum_\beta \left(A_{\alpha\beta}^{NN} f_\beta^N + A_{\alpha\beta}^{NT} f_\beta^T \right)$$

$$a_\alpha^T = b_\alpha^T + \sum_\beta \left(A_{\alpha\beta}^{TN} f_\beta^N + A_{\alpha\beta}^{TT} f_\beta^T \right) . \qquad (26)$$

The full set of consistency conditions then reads

$$a_\alpha^N \geq 0$$
$$f_\alpha^N \geq 0$$
$$a_\alpha^N f_\alpha^N = 0 \qquad (27)$$
$$|f_\alpha^T| \leq \mu f_\alpha^N$$
$$a_\alpha^T \left(|f_\alpha^T| - \mu f_\alpha^N \right) = 0 .$$

Equations (26) together with the conditions (27) can be solved with a modified Dantzig's algorithm [10] which will be discussed in the next section. Note that some of the tangential forces may be directly determined by the respective normal forces, i.e., when sliding at this contact occurs. The consistency conditions for these forces have to be fulfilled, nevertheless.

6   Dantzig's Algorithm

Contacts can be classified into breaking contacts ($a^N = 0$ thus $f^N = f^T = 0$), permanent static contacts, ($a^N = 0$ and $a^T = 0$) or permanent sliding contacts ($a^N = 0$ but $a^T \neq 0$). If we knew à priory into which category each contact belongs the contact forces could be determined by solving an inhomogeneous system of linear equations which consists of all equations for which either $a_\alpha^N = 0$ or $a_\alpha^T = 0$, with the corresponding f_α^N and f_α^T as variables. All remaining normal forces are zero, the remaining tangential forces assume their maximum values. Unfortunately the contact classification is only known if we know the contact forces as well.

We apply Dantzig's Algorithm to determine the forces together along with the corresponding contact classification. It starts with considering a certain contact, disregarding all others, i.e., their contact forces are set to zero. After having found a solution for this contact its classification is also known. The algorithm proceeds then with the next contact. Again the contact forces are determined preserving the consistency of all contacts considered before. In this process the contact classifications of the already consistent contacts may be changed if necessary. The process is repeated until the last contact has been classified.

All contacts are assigned to one of four lists:

- List NC of breaking contacts
- List C_F of permanent static contacts
- List $C^\pm$ of permanent sliding contacts. In list C^+ are all contacts where $f^T = \mu f^N$, in C^- all contacts where $f^T = -\mu f^N$.

All contacts in the above lists are considered to be consistent, i.e., they satisfy the consistency conditions. The classification is done successively for all contacts α by:

1. Check if the normal force $f_\alpha^N = 0$ satisfies the consistency conditions $a_\alpha^N \geq 0$. If this is the case the contact is consistent and belongs to NC.
2. If $a_\alpha^N < 0$ we have to increase the normal contact force f_α^N to obtain a non-negative normal acceleration. However, increasing the normal force f_α^N will change the contact accelerations of the already classified contacts as well. Since for the permanent contacts β we need $a_\beta^N = 0$ (and for the static contacts also $a_\beta^T = 0$) this would invalidate the classification of these contacts. To preserve the consistency of the classified contacts we

will have to change the contact forces of the the permanent contacts
as well. We now have to determine how much we have to change these
contact forces for a given increase s of the new force f_α^N to keep $a_\beta^N = 0$
and, if necessary $a_\beta^T = 0$. To calculate the necessary changes we formulate
a reduced set of geometry equations:

$$
\begin{aligned}
0 = a_\beta^N &= b_\beta^N + A_{\beta\gamma}^{NN,\mathrm{red}} f_\gamma^N + A_{\beta\gamma}^{NT,\mathrm{red}} f_\gamma^T + A_{\beta\alpha}^{NN} f_\alpha^N \\
0 = a_\beta^T &= b_\beta^T + A_{\beta\gamma}^{TN,\mathrm{red}} f_\gamma^N + A_{\beta\gamma}^{TT,\mathrm{red}} f_\gamma^T + A_{\beta\alpha}^{TN} f_\alpha^N .
\end{aligned}
\tag{28}
$$

The reduced set of geometry equations can be obtained from the original
geometry equations (26) disregarding all breaking contacts and replacing
$f_\beta^T = \pm\mu f_\beta^N$ for all contacts β from $C^\pm$ (sliding contacts). Since only
permanent contacts remain all contact accelerations in the reduced ge-
ometry equations are thus zero. If we now change the new contact force
$f_\alpha^N \to f_\alpha^N + s$ we have to vary the previously known contact forces $f_\beta^{\{N,T\}}$
in the reduced geometry equations by an unknown amount $\Delta f_\beta^{\{N,T\}}$ in
order to keep the contact accelerations at their required value of zero:

$$
\begin{aligned}
0 =& b_\beta^N + A_{\beta\gamma}^{NN,\mathrm{red}} \left(f_\gamma^N + \Delta f_\gamma^N \right) + A_{\beta\gamma}^{NT,\mathrm{red}} \left(f_\gamma^T + \Delta f_\gamma^T \right) \\
& + A_{\beta\alpha}^{NN} \left(f_\alpha^N + s \right) \\
0 =& b_\beta^T + A_{\beta\gamma}^{TN,\mathrm{red}} \left(f_\gamma^N + \Delta f_\gamma^N \right) + A_{\beta\gamma}^{TT,\mathrm{red}} \left(f_\gamma^T + \Delta f_\gamma^T \right) \\
& + A_{\beta\alpha}^{TN} \left(f_\alpha^N + s \right) .
\end{aligned}
\tag{29}
$$

Inserting Eq. (28) we find

$$
\begin{aligned}
0 =& A_{\beta\gamma}^{NN,\mathrm{red}} \Delta f_\gamma^N + A_{\beta\gamma}^{NT,\mathrm{red}} \Delta f_\gamma^T + A_{\beta\alpha}^{NN} s \\
0 =& A_{\beta\gamma}^{TN,\mathrm{red}} \Delta f_\gamma^N + A_{\beta\gamma}^{TT,\mathrm{red}} \Delta f_\gamma^T + A_{\beta\alpha}^{TN} s .
\end{aligned}
\tag{30}
$$

This is linear system of equations for the unknown $\Delta f_\gamma^{\{N,T\}}$ with the
step size s being a parameter. The necessary variations $\Delta f_\gamma^{\{N,T\}}$ are
proportional to the step size s, hence the solution is of the form

$$
\Delta f_\gamma^{\{N,T\}} = F_\gamma^{\{N,T\}} s ,
\tag{31}
$$

where $F_\gamma^{\{N,T\}}$ is the necessary variation if $s = 1$. By inserting the changed
values back into the original geometry equations we now know as well by
how much the accelerations of the breaking and sliding contacts change.

3. We increase the value of the new force f_α^N until either the acceleration
   $a_\alpha^N = 0$ (the contact is now consistent) or until the classification of any
   already consistent contact changes. The classification changes if either
   (a) a normal acceleration $a_\beta^N > 0$ becomes zero: The contact becomes per-
       manent and has to be moved from NC to either C_F or $C^\pm$ according
       to its present value of the corresponding tangential acceleration.

(b) a normal force $f_\beta^N > 0$ (permanent contact) becomes zero: the contact is now a breaking contact and has to be moved from C_F or $C^\pm$ to NC.

(c) a tangential acceleration $a_\beta^T \neq 0$ becomes zero and the corresponding tangential velocity is zero: the contact now is static and has to be moved from $C^\pm$ to C_F.

(d) a friction force previously of smaller magnitude than its allowed maximum reaches the maximum value $\pm\mu f_\beta^N$: the contact becomes sliding and has to be moved from C_F to $C^\pm$.

4. if we have not yet found consistent values for a_α^N and f_α^N we have to proceed with step 2.

If the contact α is permanent we have to consider the tangential component a_α^T of contact α too. The procedure is very similar to the calculation of a_α^N. The only difference is that if $a_\alpha^T > 0$ we have to *decrease* the friction force until it assumes its negative maximum value and if $a_\alpha^T < 0$ we have to increase the friction until it adopts its positive maximum value.

7 Collisions

In the framework of Rigid Body Dynamics collisions occur if two contacting particles have a negative normal relative velocity at their contact point. We can easily convince ourselves that no finite contact force can prevent a deformation of the particles. No matter how large the force is, it will always take a finite, however small, time to stop the approaching particles, hence they will deform each other. Thus, to prevent deformation of the particles an infinite repulsive force of infinitesimal duration is necessary. It turns out that the total momentum transfer Δp between the two particles is finite[2]:

$$\Delta p = \lim_{t_c \to 0} \int_0^{t_c} f\,dt = \text{finite} . \tag{32}$$

Resolving multi-particle collisions it turns out that the resulting state of the particles after the collision is not unique. This is due to the infinite stiffness of the particles, or equivalently, the infinite speed of sound in the particle material. In the limit of infinite speed of sound all information on the exact collision mechanism, e.g., the sequence of individual pair collisions, is lost, since any of the collisions is of vanishing duration. Colliding particles of finite stiffness do not exhibit this feature, since here all processes take finite time.

---

[2] Note that the duration t_c of the collision is only formally the parameter of the limit in the equation above. This limit is actually achieved by starting with deformable particles and increasing their stiffness to infinity. In this limiting process the duration of collision approaches zero.

According to these arguments the necessary information on the detailed collision mechanism is not accessible. The following set of assumptions turned out to yield realistic results, although they cannot be uniquely derived:

1. All individual pair collisions occur at once.
2. The transfer of momentum at contact points is finite. There is no momentum transfer which corresponds to attractive forces.
3. The relative velocity after a collision can never be smaller than $-\epsilon v$, where ϵ is the coefficient of restitution and v is the impact velocity at that contact ($v < 0$!).
4. If the velocity after the collision is strictly larger than $-\epsilon v$ there is no momentum transfer at this contact.

The first two assumptions have been introduced already. The remaining two assumptions deserve further discussion. Two-particle collisions can be described by means of the coefficient of restitution which relates the precollisional relative velocity v and the final velocity v' after the collision:

$$v' = -\epsilon v \ . \tag{33}$$

In the case of multi-particle collisions, however, two particles which initially rest relatively to each other may separate after a collision. Therefore, the final velocity may indeed be larger than the value $-\epsilon v$. Contrary it may no happen that two particle which collide with a finite impact velocity are at rest relative to each other afterwards. Therefore, $v' \geq -\epsilon v$.
The fourth assumption simply means, that if two particles separate from each other with higher velocity than $-\epsilon v$ their aftercollisional velocity may not be increased further by an additional momentum transfer.
For convenience we define the excess velocity Δv

$$\Delta v = v' + \epsilon v \, , \tag{34}$$

which is zero if $v' = -\epsilon v$. The source of any velocity change is a momentum transfer between the colliding particles. We can relate the excess velocities at the contacts with the momentum transfers by means of the collisional geometry equation, which is derived in a similar way as the geometry equation of the force algorithm:

$$\Delta v_\alpha = (1 + \epsilon) \, v_\alpha + A^{\text{col}}_{\alpha\beta} \Delta p_\beta \ . \tag{35}$$

Note that the v_α are relative velocities of the contact points of two contacting particles, but not the velocities of the particles themselves. In mathematical terms the above discussed assumptions read

$$\Delta v_\alpha \geq 0$$
$$\Delta p_\alpha \geq 0 \tag{36}$$
$$\Delta v_\alpha \Delta p_\alpha = 0 \ .$$

The first condition prevents the final velocity from being smaller than $-\epsilon v$. The second condition excludes attractive interaction between particles. Finally, the third condition means that there may be a finite momentum transfer only if $\Delta v = 0$. These conditions together with the geometry equation (35) form a Linear Complementarity Problem which is already familiar from Sec. 6 and can be solved by Dantzig's Algorithm.

8   Resolution of Static Indeterminacy

If the number of contacts in the system is too large, the contact forces cannot be uniquely determined by the force algorithm. This occurs, if the number of free variables in the system, i.e. the number of contact forces, is larger than the number of mechanical degrees of freedom, $3N$ in 2d or $6N$ in 3d, with N being the number of particles. However, the total forces and torques acting on the particles and, hence, their trajectories are unique. This drawback restricts the applicability of Rigid Body Dynamics for the simulation of railway ballast, since this system is of mainly static nature, the exact knowledge of the contact forces is crucial for understanding its behavior.

So far we have considered the contact forces as independent of each other. This assumption is the reason for the force indeterminacy. In realistic systems, however, the forces are not independent (Fig. 3). If we apply an external force directed to the right on the central particle, we increase the contact force with the particle to its right while at the same time decreasing the contact force with the particle to its left. In this example both contact forces in reality depend on a single parameter, which is the applied external force.
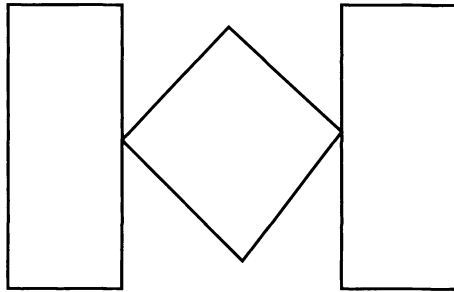
Fig. 3. The central particle is in contact with two other particles. The contact forces on both contacts are not independent of each other.

We can mimic this behavior by introducing small displacements of the particles. Each particle has a set of macroscopic coordinates r and ϕ and a set of microscopic coordinates δr and $\delta\phi$. In analogy to the vector q of all (macroscopic) coordinates of all particles, we define the vector δq of all microscopic (infinitesimal) coordinates. The contact network and the kinematic state of

the particles is determined solely by the macroscopic coordinates, hence they can be considered as the actual coordinates of the particles. The microscopic coordinates lead to a deformation of contacting particles. By definition they are of infinitesimal size, which allows us to restrict ourselves to a linear approximation in δq for the computation of the deformations. The vector ξ of all deformations in the system

$$\xi = \hat{D}\delta q \tag{37}$$

is defined by the deformation matrix $\hat{D}$ and the microscopic coordinates. The dependence of $\hat{D}$ on the geometric properties of the systems is straightforward but lengthy, therefore we will not give explicit expressions here (for the full derivation see [15]). We define a force law to relate the deformations at the contact points with the contact forces:

$$f = f(\xi)\,. \tag{38}$$

Now the contact forces are functions of the displacements δq. To calculate the forces we use the geometry equation (26) together with the consistency conditions (27). This set of equations is to be solved for the microscopic displacements δq. Hence, there are as many variables as degrees of freedom, i.e., the system is unique.

To determine the microscopic coordinates we use an overdamped relaxation method. We start with a set of inconsistent coordinates δq. Inconsistent means that the consistency conditions for the resulting forces and accelerations are not fulfilled. Now we let the system relax. This way we find new microscopic coordinates such that $\xi' = \xi - ha$, with h being the step size and a the vector of all contact accelerations. If the contact acceleration is negative the displacement will be larger, yielding a larger force to stop the approaching motion of the contacting particles. Since the adjustment step for the displacements is proportional to the acceleration the sequence of microscopic coordinates of the particles can be understood as overdamped motion. The adjusted microscopic coordinates are the solution of the linear system of equations

$$\xi - ha = \hat{D}\delta q' \tag{39}$$

or, equivalently

$$\hat{D}\left(\delta q' - \delta q\right) = -ha\,. \tag{40}$$

In cases where there are less degrees of freedom than contact accelerations the system of equations is overdetermined. There may be vectors ha which are not representable by any vector $\delta q' - \delta q$. In this case we have to project the vector $-ha$ into the image space of the operator $\hat{D}$ before solving the system of equations.

We repeat the adjustment of microscopic coordinates until the consistency conditions are met. To further improve the speed of this method we can save the microscopic coordinates which yielded a consistent system in the previous time step. If the system did not change too much this set of microscopic coordinates is very close to the new solution and we need only few iteration steps to arrive at the new solution.

This method combines the advantage of Rigid Body Dynamics, namely the ability to simulate very stiff particles, with the advantage of Molecular Dynamics, namely uniquely defined contact forces. An additional advantage of the method of small displacements is that we can now easily simulate certain degradation mechanisms. If we, for example, want to simulate the effects of abrasion of edges of the particles we can do this by gradually changing the force law (38), which describes changing edge properties of the particles.

9   Step Size Control

The integration scheme used in our implementation is a Runge-Kutta method of fourth order. During one time step there are four force computations necessary. Since we use discrete time steps we are frequently faced with the problem that after a given time step some particles do deform each other. In this case, obviously, the chosen time step was too large. This problem could be solved by predicting the time of next contact (collision) from the present state of the particles. Since the particles are subject to forces which vary in time this prediction cannot be accurate, as we approach a collision we would have to update it repeatedly. Since this prediction method is quite complicated we chose a simpler method. We advance the system by a certain time step. When determining the contacts in the next time step (see Sec. 4) we check for deformations of the particles. If any deformations occur we restore the state of the system before this time step and take a time step of half its original value. We then repeat this process, i.e. advance the system by the new time step and check for deformations. If we again encounter deformations we again divide the time step by two and repeat the computation until a state without deformations is reached. The new state of the system is accepted and the other steps of the algorithm are performed. This procedure ensures that there will be no particle deformations in the system at the beginning of any accepted time step.

Since this method can only decrease the time step, however, we need a procedure to increase the time step again[3] in order to avoid permanently slowing down the simulation. We can always increase the time step if there are no collisions in the systems for a certain time. Hence, if we successfully performed a number of time steps without encountering collisions we can increase the time step again by a factor of two. We cannot increase the time step after

---

[3] There is, of course, an upper limit to the time step which is dictated by the accuracy of the chosen integration scheme.

only one successful time step since this may lead to frequently alternating increase and decrease steps. A requirement of three successful time steps before increasing the time step has shown to yield good computational efficiency. With this choice of a step size control method we have completed the discussion of the simulation algorithm.

10   Conclusions

Although Molecular Dynamics methods have recently been applied successfully to the simulation of the dynamics of railway ballast [16] we have doubts that this technique is suitable to the simulation of almost rigid, sharply edged particles such as railway ballast, in particular if the system behavior is governed mainly by static properties of the system. Moreover, it seems to be unsuitable for the simulation of long time effects such as densification and wear of ballast.

The Rigid Body Dynamics is a method which is intended to describe the motion of systems of very stiff particles. In a natural way it avoids the problems of Molecular Dynamics simulations which have been discussed in detail in Sec. 2. These problems are unavoidable within the concept of Molecular Dynamics. Therefore we believe that Rigid Body Dynamics is much better suited to the simulation of railway ballast than Molecular Dynamics. From the presentation of the algorithm we have seen that a time step in Rigid Body Dynamics is a lot more complicated than a time step in Molecular Dynamics. Hence the according implementation requires by far more computing time than an implementation of a Molecular Dynamics code. Fortunately this disadvantage is offset by the fact that in Rigid Body Dynamics we can chose a larger time step than in Molecular Dynamics. Whereas in Molecular Dynamics the time step is determined by the critical deformation (the length on which the force changes by one unit) divided by the characteristic velocity in the system, which usually yields a very small time step (e.g. $\sim 10^{-7}$ sec), the time step in Rigid Body Dynamics is determined by the characteristic time in which the forces in the system change significantly. The latter quantity is usually much larger ($\sim 10^{-3}$ sec). Thus, the computing time requirements for both methods are comparable when simulating mainly static problems as it is the case for railway ballast.

An additional advantage of Rigid Body Dynamics is that fracture of particle can be introduced in a very natural way. When dealing with deformable particles in MD simulations the fracture of particles can lead to numerical difficulties. After the fracture of a particle, i.e., when a large particle is replaced by two or more fragments, there will be a finite time in which the system has to readjust and to reach a new stable state. There are situations when the new stable state deviates significantly from the realistic (experimental) situation. This problem is not encountered in Rigid Body Dynamics. After a fragmentation (which is algorithmically done by constructing two or more

new polygons from the previous particle simply by introducing another edge into the old particle and splitting along this line) the system remains always stable.

It is our firm believe that Rigid Body Dynamics provides a very promising alternative to Molecular Dynamics for the simulation of railway ballast.

References

1. Herrmann H.J., Hovi J.-P., Luding S. (1998) Physics of Dry Granular Media. NATO ASI Series, Dortrecht, Kluwer
2. Schäfer J., DippelS., Wolf D.E. (1996) Force schemes in simulations of granular materials. J. Physique I, 6:5
3. Brilliantov N.V. , Spahn F. , Hertzsch J.-M., Pöschel T. (1996) A model for collisions in granular gases. Phys. Rev. E 53:5382–5392
4. Lötstedt P. (1982) Mechanical systems of rigid bodies subject to unilateral constraints. SIAM J. Appl. Math. 42:281–296
5. Lötstedt P. (1981) Coulomb friction in two-dimensional rigid body systems. Z. f. Angewandte Math. und Mech. 61:605–615
6. Cottle R.W., Dantzig C.B. (1968) Complementary pivot theory of mathematical programming. Linear Algebra and its Applications, 103–125
7. Cottle R.W., Pang J.S., Stone R.E. (1992) The Linear Complementary Problem. Computer Science and Scientific Computing. Academic Press, Inc., San Diego
8. Baraff D. (1989) Analytical methods for dynamic simulation of non-penetrating rigid bodies. Computer Graphics 23:223–232
9. Baraff D. (1991) Coping with friction for non-penetrating rigid body simulation. Computer Graphics 25:31–40
10. Baraff D. (1994) Fast contact force computation for nonpenetrating rigid bodies. In: Computer Graphics Proceedings, Annual Conference Series, 23–34
11. Baraff D. (1996) Linear-time dynamics using lagrange multipliers. In: Computer Graphics Proceedings, Annual Conference Series, 137–146
12. Moreau J.J. (1994) Some numerical methods in multibody dynamics : application to granular materials. Eur. J. Mech. 13:93–114
13. Roux J.-N. (2000) On the geometric origin of mechanical properties of granular materials. Phys. Rev. E 61:6802–6836
14. Jean M. (1995) Frictional contact in collections of rigid or deformable bodies: A numerical simulation of geometrical motion. In: Salvadurai, A.P.S., Boulon M.J. (Editors) Mechanics of Geomaterial Interfaces, Amsterdam, Elsevier Science Publisher
15. Pöschel T., Schwager T. (in preparatio) Granular Dynamics. Lecture Notes in Physics, Springer, Berlin
16. Kruse H., Popp K. (2002) Model-based investigation of the dynamic behaviour of railway ballast. In: Popp K., Schiehlen W. (Editors) System Dynamics of Long-Term Behaviour of Railway Vehicles, Track and Subgrade, Lecture Notes in Applied Mechanics, Berlin, Springer

A Comparative Study of Results from Numerical Track–Subsoil Calculations

W. Ruecker[1], L. Auersch[1], M. Baessler[1],
K. Knothe[2], Y. Wu[2], U. Gerstberger[2],
K. Popp[3], H. Kruse[3],
S. Savidis[4], R. Hirschauer[4], C. Bode[4], W. Schepers[4],
G. Schmid[5], K. Friedrich[5]

[1] Federal Institute of Material Research and Testing (BAM),
Unter den Eichen 87, 12205 Berlin, Germany
[2] Technical University of Berlin, Faculty of Transport, Berlin, Germany
[3] University of Hanover, Institute of Mechanics, Hanover, Germany
[4] Technical University of Berlin, Faculty of Civil Engineering, Berlin, Germany
[5] Ruhr University Bochum, Faculty of Civil Engineering, Bochum, Germany

Abstract. The report presents the second part of the Benchmark test which was part of the DFG-priority program "System dynamics and long-term behaviour of vehicle, track and subgrade". Ballasted tracks and slab track systems on homogeneous and layered half-space are examined. Different track systems, different soil conditions, different calculation methods and a number of parameters were investigated. For example the complete or the relaxed boundary condition between the ballast and the soil was found to be of importance at low frequencies. The type and the values of the material damping of the pad element has great influence on the flexibility of a slab track. Generally it has to be noted, that the flexibility functions of all analyzed systems are described sufficiently well by all models.

1 Introduction

During the DFG-priority program "System dynamics and long-term behaviour of vehicle, track and subgrade" models for the calculation of the system "vehicle-track-subgrade" have been developed. A Benchmark test was arranged between different research partners in order to proof and to validate the models and the different calculation procedures used. This Benchmark test was coordinated by the working group of the Federal Institute of Material Research and Testing (BAM). The first comparison of calculations took place already in the second phase of the research priority program and results were presented in 1999. The main emphasis of these first examinations was the calculation of single sleepers as well as a ballasted track consisting of several sleepers and lying on a homogeneous or layered half-space. The report on hand represents the second part of the Benchmark test: the calculations of complete track systems. Ballasted tracks and slab track systems are examined on homogeneous and layered half-space.
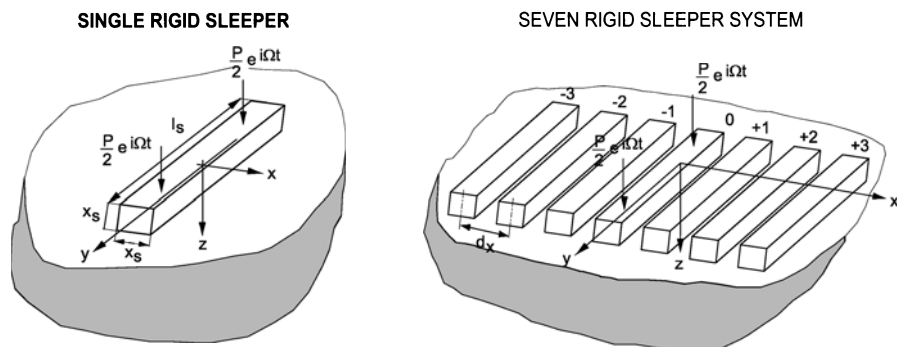
Fig. 1. Sleeper-soil systems (first part of Benchmark test)

2   Approach and Participants

The Benchmark test was activated according to the conclusion of the first phase of the priority program in 1998. Firstly different suggestions on the systems to be calculated were presented and coordinated. The definite appointments were made in the beginning of 1999. The calculations of the first two systems, single sleepers and ballasted tracks without rails (Fig. 1), were performed in the beginning of 1999. The Benchmark test was resumed in 2000 and taken to the end with the calculation of complete track systems during the year 2001.

The institutes named at the beginning of the paper contributed to the Benchmark test at least to the most important track-subsoil systems. In addition, the technical university of Hamburg-Harburg (O. v. Estorff, M. Firuziaan) and the Federal Armed Forces University Hamburg (R. Lammering, M. Plenge) were involved. The v. Estorff group joined this priority program at a later point in time and especially worked on calculations in the time domain. The Lammering group did measurements in the laboratory at systems quite similar to the ones calculated in the Benchmark test. In order to limit the scope of this paper the contributions of these participants are included in their reports in this volume.

3   Calculated Systems

3.1   System Data

The requirements of the second part of the benchmark test had already been fixed at begin of the first part. The complete description of the track subsoil systems is summarized in Tables 1 to 4. After rigid sleepers with mass -single sleepers and a seven sleeper system- were examined in the first part, a complete ballasted track with elastic sleepers, elastic pads and rails was investigated in the second part. In addition, a complete slab track system,

with underground, elastic plate, high elastic pads and rails was examined. Both a homogeneous half-space and a layer system (soft layer on stiff layer and vice versa) were regarded. The frequency domain to be examined contains frequencies of 0 to 200 Hz with a frequency step of 5 Hz.

Table 1. Parameters of the homogeneous and layered soil

Subsoil	homogeneous	soft layer	stiff layer
$v_{S,subsoil}$ in m/s	200	150	300
$\rho_{subsoil}$ in kg/m^3	1700	1700	1700
$\nu_{subsoil}$	0.25	0.25	0.25
layer thickness in m		1.0	1.0
material damping	$5 \cdot 10^{-5}$	$5 \cdot 10^{-5}$	$5 \cdot 10^{-5}$

Table 2. Parameters of the ballasted and the slab track

Slab Track		Ballast	
width in m	3.0	thickness in m	0.35
thickness protection layer in m	0.3	width (top/bottom) in m	3.60/5.60
$\rho_{concrete}$ in kg/m^3	2500	v_S in m/s	300
Young's modulus in N/m^2	$3.0 \cdot 10^{10}$	$\rho_{ballast}$ in kg/m^3	1700
thickness HGT in m	0.3	$\nu_{ballast}$	0.25
Young's modulus HGT in N/m^2	$1.2 \cdot 10^{10}$	material damping	0.03

3.2   Ballasted Track

A system with rails and thirteen elastic sleepers is calculated (Fig. 2). The ballast is modelled as an own layer. For the pad element between rail and sleepers the values of the type Zw 687 is used. The underground is a linearly

Table 3. Parameters of the rail and the sleepers

Sleeper		Rail UIC 60	
sleeper distance in m	0.6	rail distance in m	1.50
sleeper length in m	2.6	I in m^4	$3.055 \cdot 10^{-5}$
sleeper width in m	0.26	area (cross section) in m^2	$7.686 \cdot 10^{-3}$
sleeper height in m	0.2	rail mass in kg/m^3	7850
sleeper mass in kg	290	Young's modulus in N/m^2	$2.1 \cdot 10^{11}$
$\rho_{concrete}$ in kg/m^3	2145	ν_{steel}	0.3
$\nu_{sleepers}$	0.16		
Young's modulus in N/m^2	$3.0 \cdot 10^{10}$		

Table 4. Parameters of the rail pads

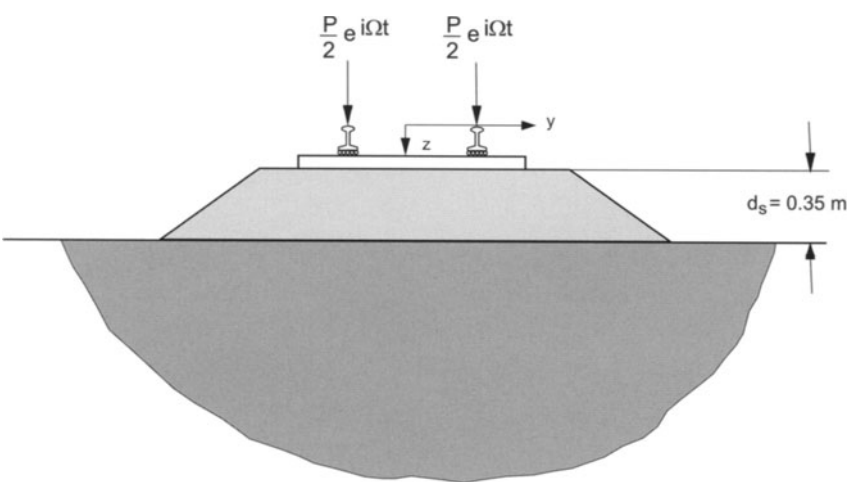| Rail Pad	Ballasted Track	Slab Track
stiffness in N/m	$6.0 \cdot 10^8$	$2.25 \cdot 10^7$
damping in Ns/m	$2.0 \cdot 10^4$	$2.0 \cdot 10^4$

Fig. 2. Geometry of the ballasted track

elastic homogeneous or a layered half-space. The track is loaded by two verti-
cal forces on both rail heads representing an axle load of 1 (N). The flexibility
functions were calculated at the top of the rails and at the sleepers.
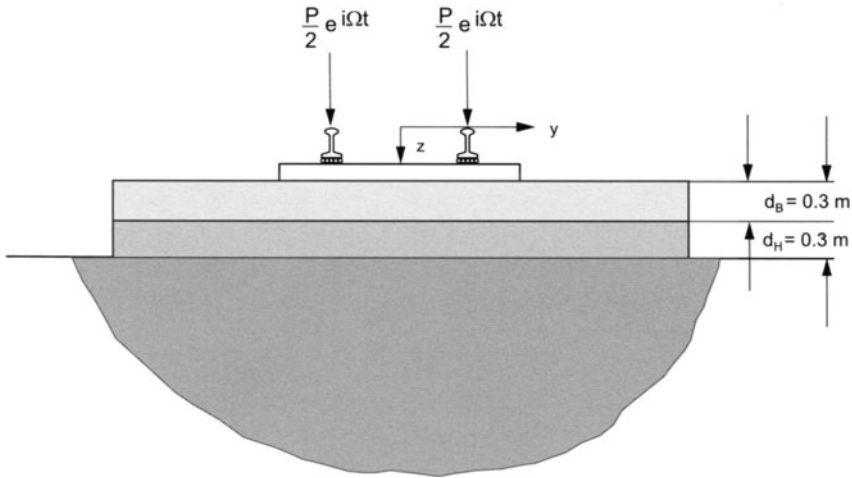

Fig. 3. Geometry of the slab track

3.3   Slab Track

As in the case of the ballasted track system, a system with thirteen elastic
sleepers is examined (Fig. 3). Under the sleepers there is a concrete plate,
which lies over a hydraulic consolidated soil (HGT). For the stiffness of the
pad element the data of the element Zw 700 have been used, which is much
softer than that used at ballasted tracks. The underground is regarded as an
elastic homogeneous or layered half-space (Fig. 4). The axle load is repre-
sented by two forces on both rail heads.

4   Calculating Procedure

The different calculation procedures developed by the participants have differ-
ent main emphases and advantages. For some procedures the best modelling
of the infinite underground is the main aim, other procedures are focused on
the possibility to model the complete system "train-track-subsoil".
To start with the underground, five different possibilities were used by the
participants, in principle. At the TU Berlin (Savidis) as well as at the BAM
the so-called thin layer method is used. This method can be used either
to formulate boundary conditions directly for the finite element range or
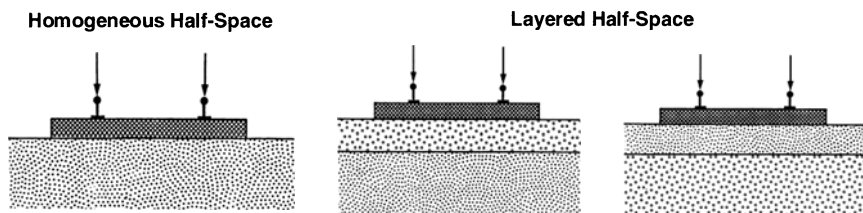else to calculate general point load solutions which are used to calculate

Homogeneous Half-Space   **Layered Half-Space**

Fig. 4. Different soil models

the flexibility matrices in a second step. For this second way complete point load solutions are calculated by BAM (with complete layer matrices of thick layers in comparison with the approximations for thin layers). The TU Berlin (Knothe) uses a solution for the homogeneous subsoil system, which is based on the exact calculation of a boundary value problem of a half-space loaded by a rectangle burden.

Another method of the calculation of flexibility matrices for the underground is the boundary element method which is used by the Ruhr University of Bochum. In this case the point load solution of the full space is used. It is the advantage of this method that one can model arbitrary layer systems, what is accompanied by a larger numeric effort.

All mentioned procedures yield flexibility matrices of the underground which must be connected with a computational model of the track. The flexibility matrices are complex and frequency dependent and very large for flexible track structures. By certain reductions (i.e rigid sleepers) or other determined simplifications (i.e. averaged deformations at the soil surfaces) it is possible to form simpler flexibility matrices. Corresponding matrices are calculated by the TU Berlin (Savidis) for the layered underground. These results were used by the TU Berlin (Knothe) and the University of Hanover in their models. The TU Berlin (Savidis) however did not use the simplified solutions, their calculation has been done with complete ground matrices.

At the TU Berlin (Savidis) the flexibility matrices of the ground are handed over to the finite element program ANSYS and connected with a FE-model of the track. The ballast, the slab track and the sleepers are modelled as an elastic continuum and the rails as elastic beams. At the generation of the coupling matrices between underground and ballast complete interaction, i.e. coupling between horizontal and vertical degrees of freedom, is assumed (non-relaxed boundary condition). For comparison purposes systems are also examined without complete interaction (relaxed boundary condition). The elastic pad element is modelled as a viscously damped spring-dashpot element. The calculation then is carried out in three steps:

- at first the point load solutions (Greens functions) of the underground are determined

- after this the dynamic flexibility matrix is built in the contact surface between ballast and underground
- the dynamic flexibility matrix of the underground is finally added to the flexibility matrix of the track and the system is solved.

The procedure of the BAM works in a similar way, however, one gets the underground stiffness and the complete solution within the same program, an extended SAP version.

Although in principle the procedures of TU (Savidis) and BAM Berlin are very similar, some differences, however, are evident:

- The Greens functions of the underground are determined with principally different procedures.
- By the generation of the stiffness matrix of the subsoil, differences exist with regard to the modelling of the contact surface between subsoil and ballast and with regard to the used integration procedures.
- Finally, there are differences in modelling the ballast, pads and rails with the finite element method with regard to the used element type as well as with regard to the degrees of freedom.

At the procedure of the Ruhr University of Bochum the underground and ballast is calculated with the procedure of boundary elements. Thereby, there is a complete contact between underground and ballast. The sleepers and rails are modelled as an elastic beam. The pad-element is regarded as a viscously damped spring-dashpot element. The boundary element flexibility matrix is coupled within the program with the finite element flexibility matrix at the contact surface of the sleepers.

TU Berlin (Knothe) and the University of Hanover used in principle an identical approach. There are, however, differences with regard to the used solution procedures. At these models the dynamic behaviour of the underground is taken into account by flexibility matrices which have been calculated in a preceding step and which assume an incomplete coupling between underground and ballast (relaxed boundary conditions). Unlike the discrete modelling with help of finite procedures the main emphasis then lies, however, on the application of analytic calculation procedure for the infinite track system. The ballast is modelled by hysteretic damped elastic bars with mass. The sleepers consist of rigid masses and the rails are Timoshenko-beams. The pad elements are taken into account as viscously damped spring-dashpot elements.

5   Results

The aim of the participants of the Benchmark test is to develop suitable calculation procedures by which the complete system "running vehicle over ballasted track or slab track system on an arbitrary underground" can be calculated sufficiently. Since the system to be looked at is extremely complex

Table 5. The different calculation methods

Element	TU Berlin (Savidis)	BAM Berlin	Ruhr-Uni Bochum	TU Berlin (Knothe)	University of Hanover
rail	elast. beam finite	elast. beam finite	elast. beam finite	elast. beam infinite	elast. beam infinite
rail pad	spring-damper	1D-bar	spring-damper	spring-damper	spring-damper
	viscous damping	hysteretic damping	viscous damping	viscous damping	viscous damping
sleeper	elast. beam	elast. beam	elast. beam	mass elem.	mass elem.
contact	non–relaxed	non–relaxed	non–relaxed	relaxed	relaxed
ballast	continuum solid FE	continuum solid FE	continuum BEM	el. bar with mass	el. bar with mass
	hysterectic damping	hysterectic damping	hysterectic damping	hysterectic damping	hysterectic damping
slab track	continuum solid FE	continuum plate and solid FE	continuum BEM		beam on visco–el. found.
subsoil	thin layer	thin and thick layer	BE-full-space	precalculated flexibility (TU Berlin)	precalculated flexibility (TU Berlin)
program	PUNCH (thin layers for subsoil)	PUNKT (thick layers for subsoil)	SSI3D	SUBTRACK	TTI
	ANSYS	BAM SAP			

and the partial systems are coupled with each other many times, simplified procedures are required. The reduction of complex models to easily comprehensible and calculable models, requires however best knowledge about the influence of the method of the modelling and about the kind of interaction of the different parameters. Therefore, in addition to the comparison of the numerical results of the different models, there is another aim of this examination: to derive notes and recommendations for a suitable and sufficient modelling of the complete system.

5.1   Ballasted Track

During the DFG-priority program different measuring campaigns were carried out by BAM with different targets (see i.e. [9]). For the comparison carried out here, measurements are selected in operational tracks although

for these measurements not all initial parameters are known. Measurements of this kind yield, however, the best summary of the real system behaviour. The result of such a measurement is presented in Fig. 5 for the ballasted track. The picture shows results of the flexibility functions of the rail, the sleeper and the subsoil. At all measurements carried out, the maximum values of the flexibility were measured in the static range. After this a steady decrease of the flexibility is generally observed, in which distinctive resonances of the system "track-subsoil" do not appear because of the high radiation damping of the subsoil. The resonance in the range around 120 Hz have to be assigned presumably to the first bending natural frequency of the sleepers. Such a behaviour, however, is not generally observed and only appears, if particularly the middle part of the sleepers is not in contact with the ballast. The flexibility function of the sleeper essentially follows the course of the flexibility of the rail but the amplitude of the flexibilities is generally smaller in dependence of the stiffness of the pad element.
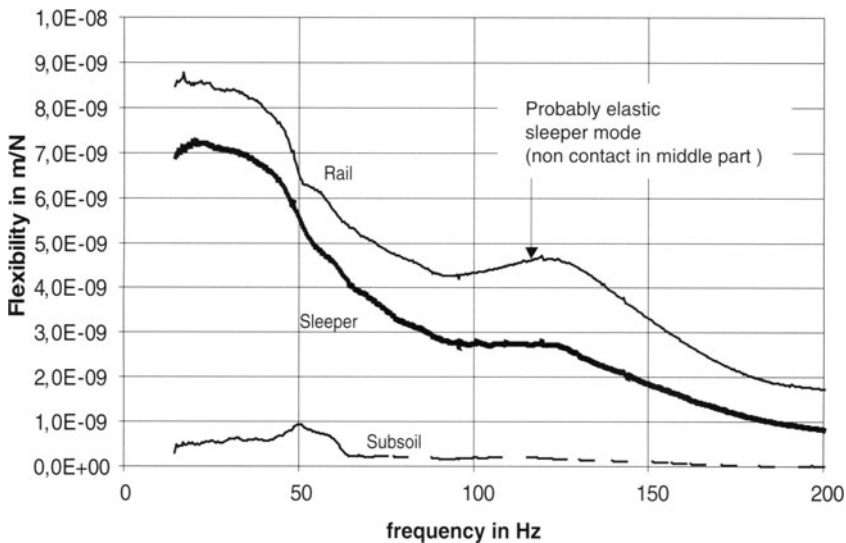

Fig. 5. Measured flexibility function of a ballasted track

The flexibility functions of the ballasted track calculated by the different participants are presented in Fig. 6 for the rail and in Fig. 7 for the sleepers. The underground is a homogeneous half-space. In agreement with the experimental results presented in Fig. 5, the numerical results show that the ballast track behaves in principle like a highly damped single degree of freedom system. The maximum values of the flexibility function are determined by the static stiffness of the track and underground. With the exception of the results of the participants who model the ballast as a block, no distinctive

resonance behaviour can be found in the frequency range examined. Larger differences arise, however, in the comparison of the absolute values of the flexibility of the single models. With regard to the model formation, the procedures of TU Berlin (Savidis) and BAM Berlin are largely comparable (see also Table 5). The comparison of the results shows that the results of the BAM and the results of the TU Berlin differ about 15%.
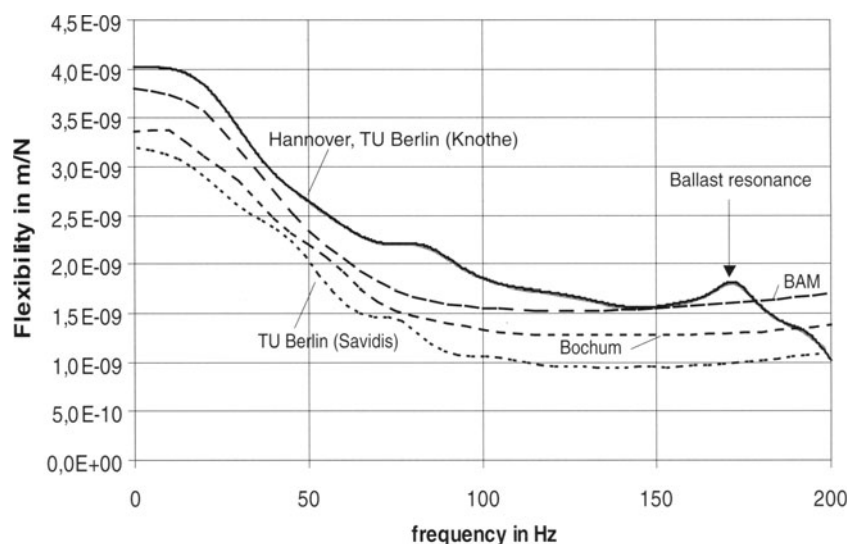

Fig. 6. Flexibility of the rail of the ballasted track on homogeneous soil

The results of the university of Hanover and TU Berlin (Knothe), which uses pre-calculated flexibility matrices for the subsoil of TU Berlin (Savidis), are in the upper end of the results. In order to come closer to the causes of these deviations and to be able to judge the sensibility of different influence parameters on the flexibility behaviour, some supplementary calculations were carried out: The influence of the sleeper number, the kind of material damping and the kind of the boundary condition between underground and ballast is investigated. The most essential results of these calculations are represented in Fig. 8.
From the flexibilities represented in Fig. 8 following results can be derived: The number of the sleepers taken into account in a track model affects only marginally the static flexibility. For frequencies higher than 20 Hz the flexibilities are almost identical. This result is in agreement with an earlier calculation in which a sufficient number of seven sleepers was stated [14]. The kind of the material damping of the ballast (viscous, hysteretic) does not have any influence on the static value, affects, however, the flexibility functions at the higher frequency range where a viscous damping leads to smaller values of the flexibility. The kind of the boundary condition between underground and
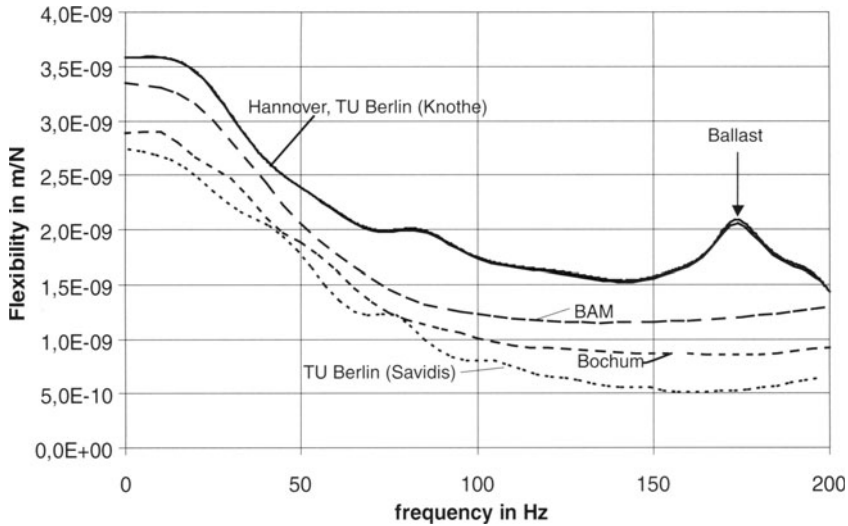
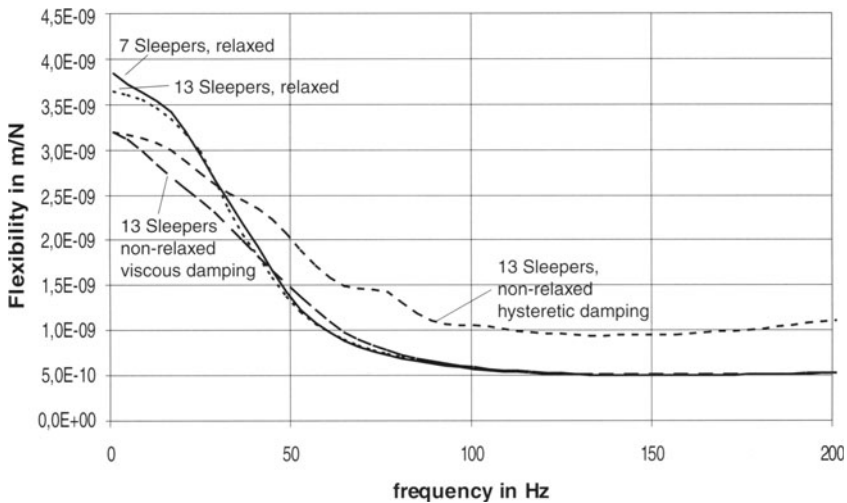**Fig. 7.** Flexibility of the sleeper of the ballasted track on homogenous soil

Fig. 8. Influence of some parameters on the flexibility of the rail

ballast proves to have a relatively strong influence on the amplitude of the flexibility functions in the static range up to frequencies of approximately 50 Hz: An incomplete coupling (relaxed boundary condition) leads to greater flexibilities by about 15%.

The following explanations for the deviations of the different models can be found on the base of the results presented in Fig. 8:

Low-frequent deviations between the analytic models and the other models

have to be ascribed mainly to the influence of the boundary condition between underground and ballast where an incomplete coupling causes larger flexibilities. The stated differences of the flexibility functions at the rail between BAM and TU Berlin (Savidis) at the higher frequency range are mainly a result of different material damping. The deviations on hand at the static and at the low-frequent range cannot be finally explained. Causes may be the different calculation of the flexibility matrices of the underground as well as the fineness of the discretization and the choice of the finite elements for ballast, pads and rail.

The results of the Ruhr University of Bochum determined with help of the boundary element method are found between the flexibilities of BAM and TU Berlin (Savidis) and the general course of the flexibility function corresponds to all other models.
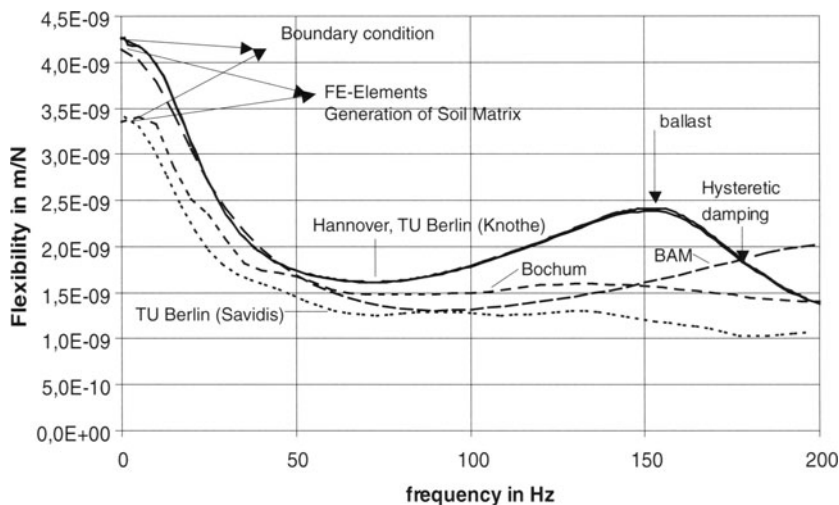

Fig. 9. Flexibility of the ballasted track on a stiff layer over soft half-space

For the layered half-spaces shown in Fig. 3 the flexibilities of the rail points of the track are represented in Fig. 9 and in Fig. 10. In the case "stiff layer over soft half-space" similar results arise as in the case of a ballasted track on a homogeneous half-space. In comparison with the results of the homogeneous half-space the static flexibilities are a little bit greater due to the influence of the softer underlying layer. After this, the decrease is stronger than in the homogeneous case. The reasons are the lower natural frequency of the complete system, the reduced radiation damping of a layer system and in the higher stiffness of the upper layer. Regarding the comparison of the different models, the same is valid as in the case of the homogeneous half-space.

Fig. 10 shows the flexibilities of the case "soft layer over stiff half-space". In difference to the flexibilities represented before, a strong system resonance
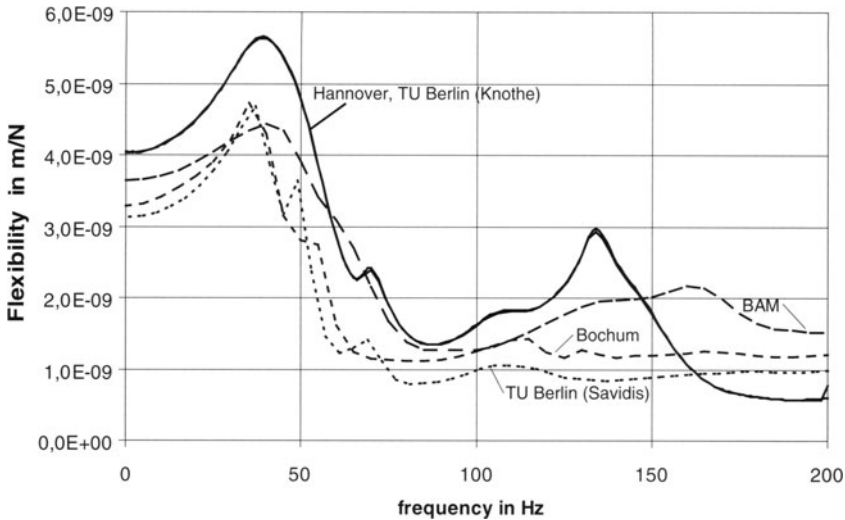
Fig. 10. Flexibility of the ballasted track on a soft layer over stiff half-space

occurs. The calculated resonance amplifications are different in the different models. The largest amplification arise in the context of the analytic models. A resonance of the ballast block which is greater than in the case of a homogeneous half-space develops in addition there. After this a good relative agreement of the results of the continuum models can be stated. The lowest amplification with an increase of the flexibility functions at the upper frequency range can be seen in the results of BAM. The reason is certainly the modelling of the pad element by a one-dimensional bar element with hysteretic damping.

In Fig. 11 the flexibilities of the rail at the adjacent sleeper no. 1 are presented. The flexibilities of this first adjacent sleeper correspond to those of the directly loaded sleeper with regard to the frequency as well as with regard to the amplitude. With regard to the frequency behaviour all models describe the flexibilities in the same way. The results of the TU Berlin (Savidis) and the Ruhr University of Bochum (Schmid) are almost identical. The differences in the results of the analytical models compared with the numerical models are caused by the different modelling of the contact surface between subsoil and ballast. The reason of the differences at the upper frequency range shown in the results of the model of BAM are due to the hysteretic damping of the pad element in comparison with the viscous damping of all other models. The deviations at the low-frequent range are created by the different formulation of the stiffness matrix of the subsoil and in the build up procedure of the track system by different finite elements.

Finally in Fig. 12 the flexibilities of the track at the third (unloaded) sleeper are given. In principle all models describe the flexibility behaviour in good agreement. The observed deviations in the static range follow the reasons
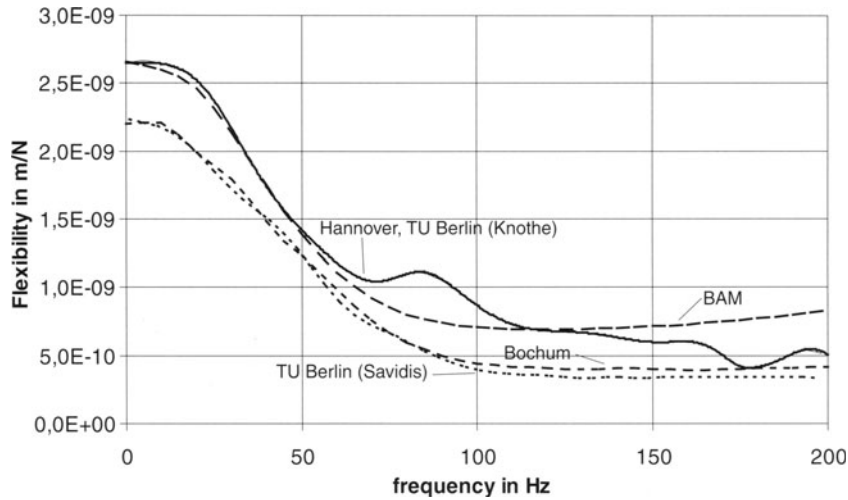
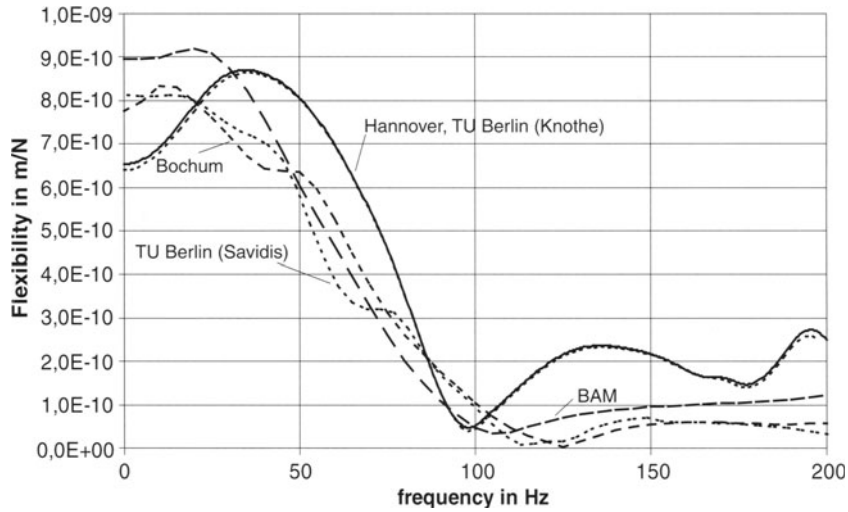**Fig. 11.** Flexibility of the rail at the first adjacent sleeper, ballasted track on homogeneous soil

Fig. 12. Flexibility of the rail at the third adjacent sleeper, ballasted track on homogeneous soil

given above. The slightly greater results of the flexibilities of the analytic models at the upper frequency range in comparison with the other models arise presumably from the missing coupling of stiffness and the missing radiation damping of the ballast modelled as a block.

5.2   Slab Track

The benchmark-test of slab track systems start with measurements made on operational tracks. In Fig. 13 flexibility functions for test points at the rail, at the sleeper or plate, at HGT and at the underground are presented. Because of the very soft pad normally built-in at this track type the largest flexibilities appear at the rail. The values of the flexibility are almost completely determined by the respective stiffness of the pad. In contrast to the ballasted track, a relatively large resonance is found at 120 Hz which arises from the rail vibrating on the pad. At higher frequencies a steady decrease of the flexibilities has to be stated. The flexibilities of the sleepers, the plate and the HGT are much smaller than those of the rail. There is hardly no difference in the magnitudes and also in the phases, which are not presented here. Therefore the conclusion can be drawn from this measurement result that the flexibilities are determined predominantly by the dynamic quantities of rail and pad for this track type. A detailed modelling of the plate, HGT and underground is of subordinate importance here.
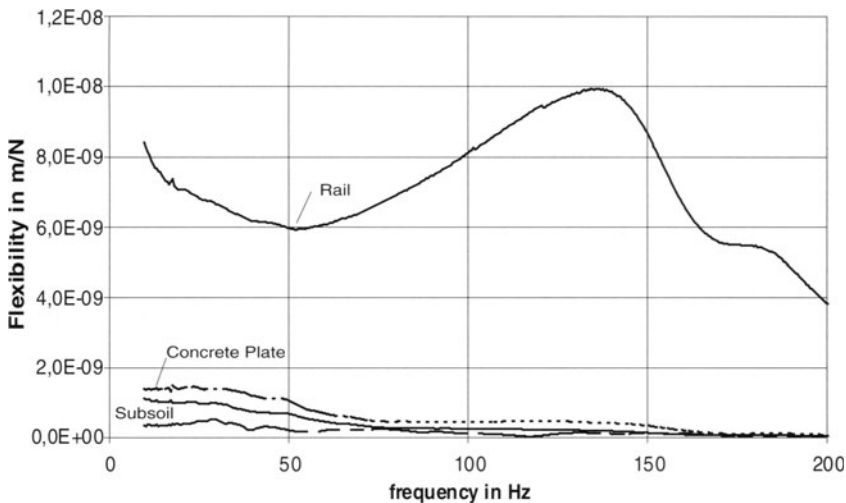

Fig. 13. Measured flexibility function of a slab track

Fig. 14 shows the flexibilities calculated with the different models of this track type. The results of all models are in very good agreement. There exist

only small differences in the static range and in the range around the rail-pad-resonance. Since the type of the modelling of the slab and the HGT is of minor importance as long as the general stiffness and masses agree, the observed small deviations have to be ascribed to the formulation of the stiffness matrix of the subsoil. The relatively small deviations in the range of the resonance arise from influences of the material damping used in the models.
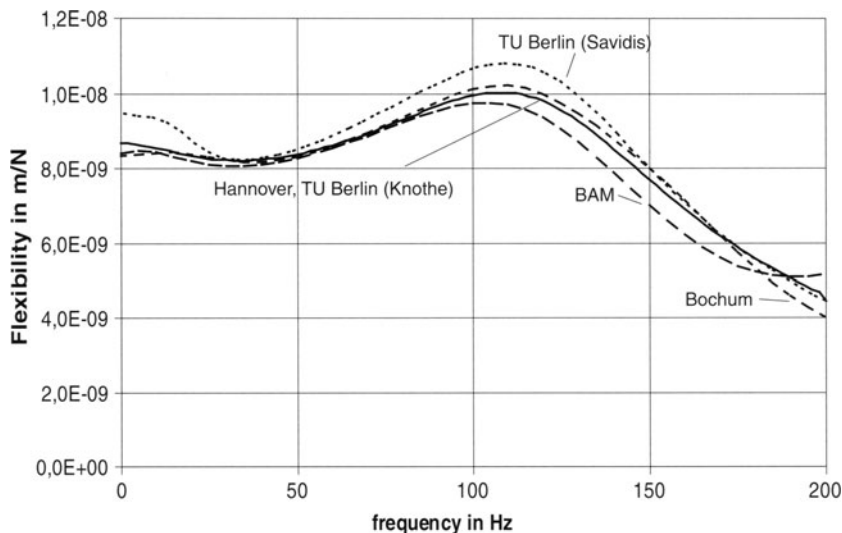

Fig. 14. Flexibility of the rail of the slab track on homogeneous soil

6   Summary

During the DFG-priority program "System dynamics and long-term behaviour of vehicle, track and subgrade" models for the calculation of the system "vehicle-track-subgrade" have been developed. A Benchmark test was arranged between different research partners in order to proof and to validate the models and the different calculation procedures used. The comparison of different tracks and modelling procedures led to the following results for the case of the ballasted track:

The number of the sleepers taken into account in a track model affects only marginally the static flexibility. For frequencies higher than 20 Hz the flexibilities are almost identical independent of the number of sleepers. The type of the material damping of the ballast (viscous, hysteretic) has no influence on the static value but affects the flexibility functions with increasing frequency where a viscous damping leads to smaller values of the flexibility functions.

The coupling boundary condition between ballast and subsoil is of relatively great importance with regard to the amplitudes of the flexibility functions up to approximately 50 Hz. In the presented results of this benchmark test an incomplete coupling of all degrees of freedom (relaxed boundary conditions) leads to about 15% greater flexibilities. Using a complete coupling the real conditions will be modelled in a better way. The modelling of the ballast as an uncoupled block element produces a resonance frequency, which is not observed by measurements. This effect can be reduced using mass proportional damping. The flexibility of the complete system is influenced mainly by the way to calculate the stiffness matrix of the subsoil, by the fineness of the discretization for the finite element part of the model as well as by the choice of the finite elements itself. All models describe the behaviour of a layered half-space sufficiently. As a result of the reduced radiation damping in the case of a layered half-space, resonance effects may be possible in block-ballast models. The flexibility functions of adjacent sleepers are described sufficient well by all models. Therefore analytic models which include the coupling through the subsoil only in an approximative way are suitable for the modelling of extended sleeper systems, too.

With respect of the obtained results for slab track systems the following conclusions can be drawn:

For all models there is a sufficient agreement between numerical and experimental results. It is not necessary to model sleepers, slab and consolidated soil separately. The stiffness of the subsoil is only slightly important to the flexibility functions of the rail and the sleepers at the frequency range up to 40 Hz. The type and the values of the material damping of the pad element has great influence on the values of the flexibility function of a slab track system.

Generally it has to be noted, that the flexibility function of all analyzed systems are described sufficiently well by all models. The deviations which could not be finally explained at the present time can be clarified with additional calculations of coordinated models.

References

1. Auersch, L. (1998) Vehicle-Track-Interaction and Soil Dynamics in: L. Palkovics (Hrsg.); The Dynamics of Vehicles on Roads and on Tracks, Supplement to Vehicle System Dynamics, Volume 29, Lisse: Swets and Zeitlinger, 553–558
2. Auersch, L. (1994) Wave propagation in layered soil: theoretical solution in wavenumber domain and experimental results of hammer and railway traffic excitation. Journal of Sound and Vibration, 173, 233–264

3. Bode, C. (2000) Numerische Verfahren zur Berechnung von Baugrund-Bauwerk-Interaktionen im Zeitbereich mittels Greenscher Funktionen fuer den Halbraum. Dissertation Technische Universitaet Berlin

4. Bode, C., Hirschauer R. and Savidis, S. (2000) Three Dimensional Time Domain Analysis of Moving Loads on Railway Tracks on Layered Soil. In: Chouw, N. and Schmid, G. (eds.) Proc. 2nd Intern. Workshop WAVE 2000, Bochum, Balkema, Rotterdam, 3–12

5. Hirschauer, R. (2001) Kopplung von Finiten Elementen mit Rand-Elementen zur Berechnung der dynamischen Baugrund-Bauwerk-Interaktion. Dissertation Technische Universitaet Berlin

6. Hubert, W., Friedrich, K., Pflanz, G., Schmid G. (2000) Frequency- and time-domain BEM analysis of rigid track and half-space with vibration barriers under investigation of causality errors for concave domains. EUROMECH Colloquium 414 'Boundary Element Methods for Soil/Structure Interaction', Catania, Italy

7. Kausel, E. and Roesset, J.M. (1981) Stiffness Matrices for Layered Soils. Bulletin of the Seismological Society of America, Vol. 71, 1743–1761.

8. Knothe, K. and Wu, Y. (1999) Rechteckfundament auf elastischem Halbraum. Bauingenieur 12, 535–541

9. Krueger, M. und Ruecker, W. (1999) Fahrzeug-Fahrweg-Wechselwirkung beim Schottergleis und bei der Festen Fahrbahn. Bericht zum DFG Schwerpunkt-projekt: Systemdynamik und Langzeitverhalten von Fahrzeug, Gleis und Untergrund, BAM Berlin

10. Kruse, H., Popp, K. (2000) The influence of wave propagation in the subsoil on the train-track dynamics. In: Chouw, N. and Schmid, G. (eds.) Proc. 2nd Intern. Workshop WAVE 2000, Bochum, Balkema, Rotterdam, 171–184

11. Kruse, H., Popp, K. (2001) A modular algorithm for linear, periodic train track models. Archive of Applied Mechanics, 473–486

12. Pflanz, G., Garcia, J., Schmid, G. (2000) Vibrations due to loads moving with sub-critical and super-critical velocities on rigid track. In: Chouw, N. and Schmid, G. (eds.) Proc. 2nd Intern. Workshop WAVE 2000, Bochum, Balkema, Rotterdam, 3–12

13. Ruecker, W. und Krueger. M. (1999) Fahrzeugschwingungen und dynamische Radaufstandskraefte eines Reisezugwagens. Bericht zum BMBF Vorhaben TV 9346, BAM, Berlin

14. Ruecker, W. (1981) Dynamische Wechselwirkung eines Schienenschwellensys-tems mit dem Untergrund. BAM-Forschungsbericht 78, Berlin

15. Ruecker, W. (1982) Dynamic interaction of a railroad bed with the subsoil. Soil Dynamics and Earthquake Eng. Conf., Southampton, Vol. 2, 435–448

16. Savidis, S., Hirschauer, R., Bode, C. and Bergmann, S. (2000) Dynamische Wechselwirkung zwischen Schienenfahrwegen und dem geschichteten Untergrund unter Beruecksichtigung von Nichtlinearitaeten. Proc. 20. Baugrundta-gung, Hannover, 285–292

17. Savidis, S., Bode, C., Hirschauer R. and Hornig, J. (1999) Dynamic Soil-Structure Interaction with Partial Uplift. Proc. 4th Europ. Conf. Struct. Dynamics - Eurodyn, Vol. II, Prag, 957–962

18. Wu, Y. and Knothe, K. (1998) Receptance behaviour of vehicle, track and subgrade. Archive of Applied Mechanics 68, 457–470.

19. Wu, Y. (1997): Semianalytische Gleismodelle zur Simulation der mittel- und hochfrequenten Fahrzeug/Fahrweg-Dynamik. VDI Fortschritt-Berichte, Reihe 12, Nr. 325, VDI-Verlag, Duesseldorf